

*Citation for published version:*

Hoffmann, JA, von Helversen, B & Rieskamp, J 2014, 'Pillars of Judgment: How Memory Abilities Affect Performance in Rule-Based and Exemplar-Based Judgments', *Journal of Experimental Psychology: General*, vol. 143, no. 6, pp. 2242-2261. <https://doi.org/10.1037/a0037989>

*DOI:*

[10.1037/a0037989](https://doi.org/10.1037/a0037989)

*Publication date:*

2014

*Document Version*

Peer reviewed version

[Link to publication](#)

©American Psychological Association, 2014. This paper is not the copy of record and may not exactly replicate the authoritative document published in the APA journal. Please do not copy or cite without author's permission. The final article is available, upon publication, at: <https://psycnet.apa.org/fulltext/2014-42221-001.html>

**University of Bath**

## **Alternative formats**

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Pillars of judgment: How memory abilities affect performance in rule-based and  
exemplar-based judgments

Janina A. Hoffmann, Bettina von Helversen, and Jörg Rieskamp

University of Basel, Department of Psychology, Basel, Switzerland

Word count: 13,129

Abstract: 151

*Author note.* Janina A. Hoffmann, Bettina von Helversen, and Jörg Rieskamp,  
Department of Psychology, University of Basel, Switzerland. This research was  
supported by two grants from the Swiss National Science Foundation  
(100014\_130192/1 and 100014\_146169/1).

Correspondence concerning this article should be addressed to Janina A. Hoffmann,  
Department of Psychology, University of Basel, Missionsstrasse 62a, 4055 Basel,  
Switzerland. Email: janina.hoffmann@unibas.ch.

### **Abstract**

Making accurate judgments is an essential skill in everyday life. Although how different memory abilities relate to categorization and judgment processes has been hotly debated, the question is far from resolved. We contribute to the solution by investigating how individual differences in memory abilities affect judgment performance in two tasks that induced rule-based or exemplar-based judgment strategies. In a study with 279 participants, we investigated how working memory and episodic memory affect judgment accuracy and strategy use. As predicted, participants switched strategies between tasks. Furthermore, structural equation modeling showed that the ability to solve rule-based tasks was predicted by working memory, whereas episodic memory predicted judgment accuracy in the exemplar-based task. Last, the probability of choosing an exemplar-based strategy was related to better episodic memory, but strategy selection was unrelated to working memory capacity. In sum, our results suggest that different memory abilities are essential for successfully adopting different judgment strategies.

*Keywords:* Judgment; working memory; episodic memory; rule-based and exemplar-based processes

“The only way to learn the rules of this Game of games is to take the usual prescribed course, which requires many years.” (Hermann Hesse)

In Hesse’s fictitious country Castalia, one of the greatest honors is to be elected Magister Ludi, the master of the glass bead game. This game integrates knowledge from all the major scholarly disciplines—from mathematics to music to philosophy—by storing this academic knowledge in the form of game symbols. During the game, these symbols are combined to form new ideas according to the grammar of the game. A challenging glass bead play thus hinges on two cornerstones of cognition: long-term memory and working memory. A glass bead player needs to store knowledge in long-term memory and retrieve this knowledge during the game; combining this knowledge requires the ability to manipulate information while keeping it activated for a short time—one key function of working memory.

Long-term memory and working memory are crucial for solving various tasks in everyday life. When shopping, for example, one must remember the items one intended to buy—a typical long-term memory task. Quickly summing up the prices in the shopping basket, by contrast, places strong demands on working memory. The ability to make accurate judgments may also hinge on basic memory processes. To judge, for instance, the attractiveness of a job offer, people may recall past work experiences from long-term memory. Alternatively, people may form an initial judgment and repeatedly update this judgment by gathering information from the job advertisement—a process that draws on key functions of working memory. These examples highlight that one can hardly think of judgments without considering memory.

Indeed, the role of memory in making judgments cannot be overstated (Weber, Goldstein, & Barlas, 1995). Consequently, the interplay of long-term and working

memory plays a major role in theories in categorization, judgment, and decision making (Ashby & O'Brien, 2005; Gigerenzer, Todd, & the ABC Research Group, 1999; Juslin, Karlsson, & Olsson, 2008; Marewski & Schooler, 2011). To what degree different categorization and judgment strategies draw on distinct memory systems has animated a particularly heated scientific debate (Ashby & O'Brien, 2005; Knowlton, 1999; Lewandowsky, 2011; Newell, Dunn, & Kalish, 2011; Nosofsky & Zaki, 1998; Smith, Patalano, & Jonides, 1998). In this vein, a growing body of research investigating the role of working memory capacity has suggested that higher capacity helps people make more accurate judgments and categorizations (Lewandowsky, 2011; Weaver & Stewart, 2012). In contrast, how long-term memory contributes to judgments and categorizations has been investigated in only a few studies (Ashby & O'Brien, 2005; Del Missier et al., 2013; Tomlinson, Marewski, & Dougherty, 2011). Furthermore, we can think of no study that considered how various memory abilities interact with different categorization or judgment strategies.

Our goal was to fill this gap and shed light on which memory abilities underlie judgments. Specifically, we investigated how individual differences in working memory and episodic memory interact with the judgment strategies people use. Focusing on two fundamental strategy types—rule based and exemplar based (Erickson & Kruschke, 1998; Juslin, Olsson, & Olsson, 2003; von Helversen & Rieskamp, 2008, 2009)—we examined how memory abilities influence the selection and execution of these judgment strategies and, ultimately, judgment performance.

We first provide an overview of memory abilities and the strategies underlying human judgments. We then explore theoretically how judgment strategies are grounded in memory and how memory abilities encourage the selection of different judgment strategies. Finally, we report an individual difference study examining how memory abilities influence judgment accuracy and strategy use.

### Memory Abilities

Memory refers to people's ability to store information. Memory research principally distinguishes long-term memory from working memory. While long-term memory stores information for minutes to years, working memory serves the purpose of manipulating information and maintaining this information in a highly active state for a short time (Atkinson & Shiffrin, 1968). In recent theories, working memory is said to consist of activated representations in long-term memory (Oberauer, 2009; Unsworth & Engle, 2007). Evidence from individual difference studies suggests that working memory correlates with performance in long-term memory tasks (Del Missier et al., 2013; Mogle, Lovett, Stawski, & Sliwinski, 2008; Unsworth, 2010). Specifically, working memory may control encoding into and strategic retrieval from long-term memory (Baddeley, Lewis, Eldridge, & Thomson, 1984; Craik, Govoni, Naveh-Benjamin, & Anderson, 1996; Rosen & Engle, 1997; Unsworth, Brewer, & Spillers, 2013).

Furthermore, memory research has distinguished implicit from explicit long-term memory (we use the term *episodic memory* to refer to explicit long-term memory of specific events). Whereas explicit memory measures reflect conscious recollection of facts or episodes, in implicit memory tests previous experiences facilitate performance without necessarily requiring their conscious recollection (Roediger, 1990; Squire & Zola, 1996). Countless studies have shown dissociations between implicit and explicit memory tests and these dissociations have often been taken as evidence of two distinct memory systems (Squire & Zola, 1996). For instance, implicit memory measures, such as word stem completion, are not correlated with episodic memory measures, such as cued recall (Bruss & Mitchell, 2009; Fleischman, Wilson, Gabrieli, Bienias, & Bennett, 2004; Perruchet & Beaveux, 1989). At the same time, the idea that there exist distinct systems has been repeatedly challenged (e.g., Berry,

Shanks, Speekenbrink, & Henson, 2012; Dew & Cabeza, 2011; Roediger, 1990). Recently, for instance, Berry et al. (2012) suggested that a single process model accommodates performance differences between recognition and implicit repetition priming. In addition, several studies raised methodological concerns about the reliability of implicit memory measures (Buchner & Brandt, 2003; Buchner & Wippich, 2000; Meier & Perrig, 2000). It remains an open question if memory can best be understood as two distinct systems.

### **Judgment Strategies**

People make judgments every day ranging from estimating the probability of rainfall to judging the attractiveness of a job. Making such judgments requires inferring a continuous criterion, for instance, job attractiveness, from a number of attributes of this object (i.e., the cues), such as the annual salary or the task demands. People may rely on two different types of judgment strategies: rule based and exemplar based (Erickson & Kruschke, 1998; Juslin et al., 2003; von Helversen & Rieskamp, 2008, 2009).

Rule-based strategies assume that people form hypotheses about the relationship between the cues and the criterion and apply this knowledge to make a judgment (Brehmer, 1994; Juslin et al., 2008). Rule-based judgment strategies have been predominantly captured with linear, additive models (Cooksey, 1996) or cue abstraction models (Juslin et al., 2003). Linear models describe people's judgments in a variety of tasks ranging from personal selection (Graves & Karren, 1992) to medical diagnoses (Wigton, 1996) and have been found to match people's explicit judgment rules (Einhorn, Kleinmuntz, & Kleinmuntz, 1979; Lagnado, Newell, Kahan, & Shanks, 2006). Based on the lens model (Brunswik, 1956), linear models assume that people explicitly abstract a weight for each cue and then combine the weighted cue values additively (Einhorn et al., 1979, Juslin et al., 2003). For instance, when judging

the attractiveness of a job offer, people first determine how much they value salary and task demands. Then they weight the annual salary and task demands of the job by their importance and combine this knowledge by adding the weighted cue values.

Exemplar-based judgment strategies, by contrast, rely on the retrieval of past experiences from long-term memory. Exemplar-based strategies assume that previously encountered objects are stored in memory along with their criterion values (Juslin et al., 2003, 2008). To judge the new object (the probe), all previously encountered objects (exemplars) and the associated criterion values are retrieved from memory. For instance, when judging the attractiveness of a new job, people may think about all past jobs they have held. The more similar a retrieved exemplar is to the probe, the more it influences the final judgment. Accordingly, previous jobs with task demands similar to the job offer influence the attractiveness rating more than unrelated work experiences. Thus, exemplar-based strategies imply that people store concrete instances without abstracting any knowledge and engage in an associative similarity-based process during retrieval.

In sum, rule-based and exemplar-based strategies differ in their assumptions about the cognitive processes underlying judgments (Hahn & Chater, 1998; Juslin et al., 2003). Whereas rule-based strategies use abstracted knowledge about the world to reason about new instances, similarity-based or exemplar-based strategies rely on similarity to past instances. Research suggests that people use both types of strategy, with strategy selection depending on task characteristics (Juslin et al., 2003, 2008; Karlsson, Juslin, & Olsson, 2007; Platzer & Bröder, 2013; von Helversen, Karlsson, Mata, & Wilke, 2013; von Helversen & Rieskamp, 2009) and individual differences (Mata, von Helversen, Karlsson, & Cüpper, 2012; von Helversen, Mata, & Olsson, 2010): Specifically, when people perform the judgment task over trials and receive feedback about the correct criterion, they rely more on cue abstraction strategies if the



criterion is a linear additive function of the cues (in a “linear task”). However, people shift to exemplar-based strategies in “multiplicative tasks” where the judgment criterion is a nonlinear function of the cues (Hoffmann, von Helversen, & Rieskamp, 2013; Juslin et al., 2008). This shift presumably takes place because the cue abstraction strategy does not allow accurate judgments in nonlinear environments (Juslin et al., 2008; von Helversen & Rieskamp, 2009). In the following section we review theoretical and empirical work on how the cognitive processes underlying rule-based and exemplar-based strategies map onto different memory abilities.

### **Linking Judgment Strategies and Memory Abilities**

In general, memory abilities can influence two aspects of strategy use: execution (i.e., the ability to execute a strategy correctly) and selection (Beach & Mitchell, 1978; Lemaire & Siegler, 1995; Mata, Pachur, et al., 2012). Regarding execution, better episodic memory can enhance exemplar retrieval and thus lead to more accurate exemplar-based judgments. Regarding selection, memory abilities can boost either the ability to choose the more accurate strategy or the preference for one strategy (Beach & Mitchell, 1978). We first address how executing rule-based and exemplar-based strategies is related to working memory, episodic memory, and implicit memory and thereafter address strategy selection.

### **The Influence of Memory Abilities on Strategy Execution**

**Rule-based strategies.** Solving a rule-based categorization or judgment task has often been equated with logical reasoning (Ashby & O’Brien, 2005) or problem solving (Juslin et al., 2008). Like reasoning or problem solving, rule-based strategies are thought to involve a serial, controlled hypothesis-testing process and, in turn, working memory (Ashby & O’Brien, 2005; Brehmer, 1994; Juslin et al., 2003, 2008). Working memory may be required by two aspects of the rule-based process: rule abstraction and rule execution.

Rule abstraction requires learning the cue weight, the weight that should be given to a specific cue. To achieve this one could compare two successively presented objects, relate the difference in judgment criteria to the difference in cue values, and then update the cue weights (Juslin et al., 2008; Pachur & Olsson, 2012). It has been argued that this comparison process is inherently sequential and capacity constrained and—as a consequence—restricts people to abstract linear, additive rules (Juslin et al., 2008). In addition, comparing two objects likely taxes working memory, because it involves storing information about the two judgment objects for a short time and manipulating this information, key functions of working memory (Atkinson & Shiffrin, 1968; Baddeley & Hitch, 1974). Overall, recent research supports the idea that learning rules hinges on working memory. Learning one-dimensional categorization rules, for instance, is impaired by a concurrent verbal task (Filoteo, Lauritzen, & Maddox, 2010; Zeithamova & Maddox, 2006, 2007). In a similar vein, cognitive load studies suggest that people abandon cue abstraction strategies more frequently under cognitive load than without load (Hoffmann et al., 2013). Finally, learning a judgment task is easier if the sequence reduces working memory demands by facilitating a direct comparison of cue values and judgment criteria (Helsdingen, Van Gog, & Van Merriënboer, 2011; Juslin et al., 2008).

Applying a learned rule may also involve working memory, such as mental updating and inhibition (Miyake et al., 2000; Oberauer, 2009). When making a judgment people may start with an initial estimate that is updated with each new piece of evidence (Hogarth & Einhorn, 1992; Juslin et al., 2008)—a process that requires keeping the past estimate in mind and manipulating it mentally. Furthermore, rule application requires inhibiting information, because people need to focus attention on the relevant cues and ignore the irrelevant ones. In line with this idea, Del Missier et al. (2013) found that correctly applying decision rules was related to working memory

capacity. Specifically, rule application involved inhibiting irrelevant information and updating information in working memory (Del Missier, Mäntylä, & Bruine de Bruin, 2010, 2012).

Long-term memory may be less important than working memory for making rule-based judgments. Once a rule has been established, only the cue weights need to be retrieved from long-term memory (Bruner, Goodnow, & Austin, 1956). Because previously encountered objects can be forgotten (von Helversen & Rieskamp, 2008), episodic memory should marginally influence rule execution.

**Exemplar-based strategies.** Exemplar-based strategies assume that judgments are based on the similarity to previously encountered exemplars (Juslin et al., 2003; Medin & Schaffer, 1978; Nosofsky, 1988), suggesting that executing exemplar-based strategies should be linked to episodic memory (Hintzman, 1986, 1988; Nosofsky, 1988). Two major processes in episodic memory may contribute to successfully adopting exemplar-based strategies: encoding into and retrieval from episodic memory (Estes, 1986; Shiffrin & Atkinson, 1969).

Before any information can be recalled from memory, it is necessary to form a memory representation (i.e., to encode) and store this information (Estes, 1986). Like episodic trace models of episodic memory, for instance, MINERVA 2 (Hintzman, 1984, 1986) or MINERVA-DM (Dougherty, Gettys, & Ogden, 1999), most exemplar-based models assume that each presentation of an exemplar is stored as a separate memory trace (Estes, 1986; Nosofsky, 1988). Accordingly, the more often an object is presented, the more often it is encoded and the more likely is its subsequent retrieval. Likewise, elaboration, adding information to the memory trace, or spacing exemplar presentations across time intervals can deepen encoding (Brown & Craik, 2000; Martin, 1968). Beyond encoding, an exemplar-based strategy also requires accurately retrieving exemplars from episodic memory. Retrieval may fail because the probe's

features—serving as retrieval cues—do not activate memory traces for stored exemplars or because past exemplars can no longer be discriminated (Medin & Schaffer, 1978).

Although theoretical accounts suggest strong links between episodic memory and exemplar-based strategies, empirical evidence for the relationship is still scarce (Ashby & O'Brien, 2005). Nevertheless, researchers have shown that the instruction to learn all exemplars by heart helps learning in judgment tasks solvable by exemplar strategies (Olsson, Enkvist, & Juslin, 2006). If single exemplars have to be memorized to solve a categorization task, these exemplars are recognized more easily in a subsequent recognition test (Davis, Love, & Preston, 2012; Palmeri & Nosofsky, 1995; Sakamoto & Love, 2004). In contrast, if people cannot identify past exemplars, they are less inclined to follow exemplar-based strategies (Rouder & Ratcliff, 2004). Furthermore, similar to spacing effects in memory (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006), spacing exemplar repetitions helps when solving exemplar-based tasks (McDaniel, Fadler, & Pashler, 2013).

Neuropsychological work has challenged the view that similarity-based category learning depends solely on episodic memory (Knowlton, 1999; Smith, 2008). The multiple-systems view (Ashby & O'Brien, 2005; Smith & Grossman, 2008) proposes instead that implicit perceptual memory underlies categorizations based upon the similarity to one prototype. Amnesiac patients, for instance, classify new dot patterns with the same accuracy as healthy controls in prototype distortion tasks but are less accurate at recognizing previously encountered patterns (Knowlton & Squire, 1993). It is possible that exemplar-based judgments might also rely on implicit perceptual memory. Although exemplar-based judgments require learning more than one exemplar to elicit different judgments, memory research indicates that more than one exemplar can be strengthened by past experience (Chiu, 2000; Musen &

Treisman, 1990). On the other hand, a single exemplar model can explain dissociations between categorization and recognition (Nosofsky & Zaki, 1998). Furthermore, even without previous training, healthy participants can achieve the same performance as trained amnesiac patients in prototype classification while showing chance recognition (Palmeri & Flanery, 1999). Thus, if implicit perceptual memory is related to exemplar-based judgments remains an open question (we discuss implicit procedural memory in the General Discussion).

Working memory could also be involved in learning exemplar-based judgments. Lewandowsky (2011), for instance, argued that every recollection-based long-term memory task should be related to working memory capacity. Underpinning his argument, working memory has been found to support encoding and retrieval in episodic memory (Baddeley et al., 1984; Craik et al., 1996; Rosen & Engle, 1997; Unsworth et al., 2013). Exemplar retrieval may also involve a deliberative search in long-term memory (Juslin et al., 2008; Karlsson, Juslin, & Olsson, 2008). Indeed, research suggests that working memory load not only harms rule-based strategies but also disturbs retrieving past exemplars when judging new objects (Juslin et al., 2008). Furthermore, learning to solve rule-based and exemplar-based categorization tasks is facilitated by high working memory capacity (Craig & Lewandowsky, 2012; Lewandowsky, 2011; Lewandowsky, Yang, Newell, & Kalish, 2012). Therefore, working memory capacity should—in general—promote executing exemplar-based judgment strategies. However, if working memory promotes exemplar-based processing by enhancing episodic memory, episodic memory will serve as a mediator between working memory capacity and exemplar-based judgments, and hence, working memory capacity should lose importance for predicting exemplar-based judgments.

### **The Influence of Memory on Strategy Selection**

When choosing a strategy, people may learn the benefits and costs associated with each strategy (Beach & Mitchell, 1978; Payne, Bettman, & Johnson, 1993; Rieskamp, 2006; Rieskamp & Otto, 2006). Hence, memory abilities could determine the preference for employing a specific strategy. Thus, people with good episodic memory may select an exemplar-based strategy more often, whereas those with bad episodic memory may avoid that type of strategy. Along these lines, researchers have found that older adults avoid exemplar-based strategies—possibly because they place high demands on episodic memory (Mata, von Helversen et al., 2012). Similarly, high working memory capacity may facilitate using rules and thus encourage rule-based processing.

There is also reason to believe that memory abilities differentially affect selecting a rule- or exemplar-based strategy. When learning to make judgments, people seem to start with a rule and switch to an exemplar-based strategy only if the rule fails (Juslin et al., 2008; Nosofsky, Palmeri, & McKinley, 1994). If rule-based strategies serve as a default option, memory abilities such as high working memory capacity may not be required to select such a strategy, only to execute it successfully.

Memory abilities may influence the general ability to choose strategies adaptively (Mata, Pachur et al., 2012). Bröder (2003) found that intelligent participants were more likely to select a strategy that ignores information if this strategy performed well. People with higher working memory capacity do not simply prefer rule-based strategies in categorization; instead they seem to select the more appropriate strategy for the task at hand (Lewandowsky et al., 2012). People with high working memory capacity may not only apply rules more accurately but also be faster in detecting when rules cannot properly solve a judgment task, prompting a shift to exemplar-based strategies.

In contrast, some studies raised the possibility that working memory capacity is not involved in strategy selection. For instance, Craig and Lewandowsky (2012) found in a 5-4 categorization task that working memory capacity did not predict choice between one-dimensional rule-based strategies and exemplar-based strategies. Similarly, if the success rates of different strategies change over time, adaptation to these changes is not predicted by working memory capacity but by awareness of those changes (Schunn, Lovett, & Reder, 2001). These results suggest it is possible that neither the decision to store exemplars nor the decision to apply a rule hinges upon working memory capacity.

### **Predictions for Judgment Performance and Strategy Selection**

To predict how memory abilities are related to judgment, it is necessary to consider the task. Judgment research suggests that people prefer rule-based strategies in linear tasks but switch to exemplar-based strategies in multiplicative tasks (Hoffmann et al., 2013; Juslin et al., 2008). Thus, memory abilities should differentially affect performance in linear and multiplicative tasks.

Specifically, low working memory capacity should harm the execution of cue abstraction strategies, because incorrect cue weights are learned or because applying the rule is disrupted. In contrast, episodic memory should only marginally influence the execution of cue abstraction strategies above working memory. Consequently, higher working memory capacity but not better episodic memory should predict more accurate judgments in linear tasks. Successfully executing an exemplar-based strategy, in contrast, hinges on encoding into and retrieval from episodic memory so that better episodic memory—and possibly implicit perceptual memory—should improve judgment accuracy in multiplicative tasks. Working memory should not affect performance in a multiplicative task above episodic memory. Regarding strategy selection, working memory capacity may help people detect and choose the more

appropriate strategy. Episodic memory may make it more likely that people rely on retrieval of past exemplars in multiplicative tasks.

### **The Present Study**

We examined how memory abilities relate to judgment performance in a linear, additive task and a multiplicative task. Additionally, we measured working memory, episodic memory, and implicit memory with three different tests each. We selected the memory tests so that variance stemming from material or task-specific effects was reduced, allowing us to measure relatively pure latent abilities (Miyake et al., 2000). These tests included different types of material (verbal, spatial, numeric) and different types of tasks (e.g., recognition, cued recall, and free recall for episodic memory).

### **Participants**

Two hundred and seventy-nine participants (147 female, 132 male,  $M_{\text{Age}} = 24.0$  years,  $SD_{\text{Age}} = 6.0$ ) were recruited at the University of Basel. Participants received an hourly fee for their participation (CHF 20, approx. U.S. \$22) and could earn an additional bonus in the judgment tasks ( $M = \text{CHF } 10.3$ ,  $SD = 2.5$ ). Overall, it took participants about 5 hr to complete the study, including a half-hour break.

### **Automated Working Memory Span Tasks**

Working memory span tasks were designed to measure both storage and processing of information in working memory (Redick et al., 2012) by letting participants process one set of stimuli while remembering another set. For instance, in each trial of the operation span task, participants first see a simple equation. After they solve the equation and give the answer, they see the first letter that has to be remembered. Subsequently, a second equation is presented and solved and then the next letter that has to be remembered is presented. Solving of equations and presentation of letters continues until a certain number of letters (the set size) has been presented. At the end of each trial, participants have to recall the letters in order of



their appearance. Trials with different set sizes are randomly interspersed, with each set size repeated three times.

We used three well-known span tasks (Unsworth, McMillan, Brewer, & Spillers, 2012; Unsworth, Redick, Heitz, Broadway, & Engle, 2009): the reading, the operation, and the symmetry span. The tasks were taken from Unsworth et al. (2009) and translated into German. We measured working memory capacity using the partial credit score (the sum of all items recalled in the correct position over all trials) as the dependent variable (Conway et al., 2005).

**Operation span.** Participants were asked to solve equations while remembering letters. Set size varied from 3 to 7 so that partial credit scores could range from 0 to 75.

**Reading span.** In the reading span participants judged the plausibility of a sentence while remembering letters.<sup>1</sup> Set size varied from 3 to 7 so that partial credit scores could range from 0 to 75.

**Symmetry span.** Participants judged the symmetry of a chessboard picture while remembering the positions of squares in a  $4 \times 4$  matrix. In each trial, participants first saw a chessboard picture and were asked to judge its symmetry. Afterward, one square in the  $4 \times 4$  matrix was highlighted and participants were asked to remember its position. After the set size had been reached, participants recalled the positions of the squares by clicking on the squares in order of their appearance. Set size varied from 2 to 5 so that partial credit scores could range from 0 to 42.

### **Episodic Memory Tasks**

We measured episodic memory with three tasks: free recall of pictures, cued recall of numbers, and recognition of verbs.

**Picture free recall.** We selected 20 pictures with high ratings on imagery and concreteness from a picture database (Rossion & Pourtois, 2004). Each picture was

presented for 3 s and participants had to remember them. After a 2-min retention interval participants recalled the pictures by naming them. Performance was measured as the percentage of correctly recalled pictures.

**Cued number recall.** We assessed cued number recall with a computerized version of the cued number recall task from the Berliner Intelligenzstruktur-Test Form 4 (BIS 4; Jäger, Süß, & Beauducel, 1997). Fifteen pairs of a two- and a three-digit number were first presented for 10 s each. After a 2-min retention interval, participants saw the cued number pair as well as 4 three-digit number distractors and had to indicate which three-digit number was initially presented together with the two-digit number. Performance was measured as the percentage of correctly recalled three-digit numbers.

**Verb recognition.** We selected 40 verbs with five to seven letters that were rated high on imagery and concreteness from the Hager and Hasselhorn (1994) database. Twenty verbs were assigned to a list of old items and 20 to a list of new items with the lists matched on word length, imagery, and concreteness. In the study phase, participants learned the old verbs for 3 s each. After a retention interval of 2 min, participants indicated whether they recognized the 40 verbs from the study phase by classifying them as *old* or *new*. Performance was measured as the percentage of verbs correctly classified as old or new.

### **Implicit Perceptual Memory Tasks**

Researchers have questioned the reliability of implicit memory measures (Buchner & Brandt, 2003; Buchner & Wippich, 2000; Meier & Perrig, 2000). To increase reliability, we used performance tests that always had a correct solution (instead of association tests such as word stem completion; Buchner & Brandt, 2003). Our participants solved three implicit perceptual memory tests: speeded presentation

of line drawings, identification of sounds presented in noise, and identification of degraded nouns.

We measured performance as the difference in median reaction times between old and new items, including correct and incorrect answers. Negative reaction time differences indicate that participants responded faster to old items than to new items, showing an implicit memory effect, a facilitation because of prior experience.

**Speeded presentation of line drawings.** We based our speeded presentation task on an experiment by Musen and Treisman (1990). We randomly created 500 line drawings, from which we excluded duplicates and simple forms, such as arrows. For the implicit memory test, we randomly selected 40 line drawings—20 old and 20 new—with the restriction to have at most two lines in common. To determine the presentation threshold we used 40 line drawings that differed from the drawings in the implicit memory test in at least two lines.

Using a threshold procedure we first determined the presentation length at which participants could correctly reproduce half of the line drawings. Starting with a presentation length of 400 frames (approx. 1200 ms), participants were asked to retrace the line drawing on a mask composed of all possible lines. Participants were required to draw all five lines and told to guess if they could not remember a line. Presentation length decreased by 100 frames (300 ms) after each correct reproduction and increased by 100 frames after each incorrect one. Step size decreased 10 frames (30 ms) after five turning points (turning point refers to a switch between decreases and increases in presentation length).

In the subsequent study phase, participants had to click as fast as possible on all lines of the 20 old items. Participants retraced all old items twice. After a 2-min retention interval, participants again completed a speeded reproduction task. The presentation length was set to the presentation length at the end of the threshold

procedure. Participants had to redraw the briefly presented old and new line drawings. Performance was measured as the difference in median reaction times between old and new line drawings.

**Identification of degraded nouns.** Forty 5- to 7-letter nouns with high imagery and concreteness were selected from the Hager and Hasselhorn (1994) database. Nouns highly similar in spelling were excluded. The nouns were alphabetically sorted and 20 items with the same initials were randomly included in the old and new item list. To present the nouns in a degraded fashion, we superimposed an  $8 \times 2$  chessboard mask over each noun. Nine of the 16 squares were randomly turned black to render noun identification difficult.<sup>2</sup>

In the study phase, participants had to count the vowels in 20 nouns, with German umlauts counting as two vowels. Each noun was presented for 3 s on screen. After a 2-min retention interval, participants were asked to correctly identify 40 degraded nouns by typing in the noun names and confirming their response by pressing “Enter.” Half of the nouns were *old* (i.e., had already been presented in the study phase). Performance was measured as the difference in median reaction times between old and new degraded nouns with reaction time operationalized as time to last key press.

**Sound identification in noise.** We selected 40 sounds from the Database for Environmental Sound Research and Application (Gygi & Shafiro, 2010) with a length between 0.55 and 3.54 s. All sounds were equalized for RMS (root mean squared) loudness, so that mean RMS loudness was 60 dB. The median spectral centroid—a measure of central tendency that characterizes the frequency spectrum and is correlated with subjective brightness of a sound (Schubert & Wolfe, 2006)—ranged from 262 to 5,507 Hz. For the sound identification task, the sounds were embedded in

5 s of white noise with a signal-to-noise ratio of -15 dB. Each sound started 0.5 s after stimulus onset.

In the study phase, participants were asked to indicate whether the 20 old sounds had a higher or lower pitch than their own voice. After a 2-min retention interval, participants listened to 20 old sounds from the study phase and 20 new sounds, all embedded in noise.<sup>3</sup> After each sound, participants were shown the name of the sound as well as the names of two other sounds that never appeared in the study and had to indicate which of the sounds they had listened to. Performance was measured as the difference in median reaction times between old and new sounds.

### **Judgment Tasks**

Participants solved both a linear and a multiplicative task, taken from Hoffmann, von Helversen, and Rieskamp (2014). In both tasks, participants had to judge a continuous criterion ranging from 0 to 50 based on four cues varying on a continuous scale from 0 to 5. In the linear task, the criterion  $y$  was a linear, additive function of the cues:

$$y = 4c_1 + 3c_2 + 2c_3 + c_4, \quad (1)$$

where  $c_1$  reflects the most important cue according to its cue weight. Each cue value varied between 0 and 5. In the multiplicative task the function included a multiplicative combination of the cues:

$$y = \frac{4c_1 + 3c_2 + 2c_3 + c_4 + 2c_1c_2c_3 + c_2c_3c_4}{8.5} \quad (2)$$

We used two different cover stories for the linear and the multiplicative task. In the linear task, participants judged whether a comic figure was a good or bad catcher of small creatures. In the multiplicative task, participants estimated the toxicity of a bug. The stimuli consisted of pictures of either bugs or comic figures. These bugs and comic figures varied on four cues. The bugs varied on the length of their legs, their

antennae, and their wings, and the number of spots on their back. The comic figures had different sizes of ears and nose and a different number of hairs and stripes on their shirt.

From all possible items, we selected a subset for the training and the validation set (see Hoffmann et al., 2014): The training set was used in the training phase to allow participants to learn how to solve the tasks. The validation set was employed in the test phase to identify the judgment strategy that described participants best. The cue values were sampled from a uniform distribution for each cue. In the linear task, a linear rule-based judgment strategy made more accurate predictions for validation items and should, consequently, lead to a better performance than an exemplar-based strategy in the test phase; in the multiplicative task, however, an exemplar-based strategy not only fitted the training set better than a linear rule-based judgment strategy but also made more accurate predictions for validation items. Hence, an exemplar-based strategy should lead to a better performance in the multiplicative task. Additionally, the rule-based and the exemplar-based strategy predicted different responses on the validation items. Correlations between the cues were low in the training set, ranging from  $r = -.17$  to  $r = .11$ .

Table 1 illustrates the task structure: The cues  $c_1$  to  $c_4$  could be used to predict the correct criterion value. These cues were randomly assigned to the pictorial cues (e.g., ears or nose). Higher cue values, however, were always associated with more salient pictorial features. For instance, a cue value of zero corresponded to a bug without spots on the back and a cue value of five to a bug with five spots on its back.

Both tasks consisted of a training and a test phase. During training, participants learned to estimate the criterion values for 25 items from the training set. In each trial, participants first saw a picture of a bug or a comic figure and were asked to estimate its criterion value. Afterward they received feedback about the correct value, their own

estimate, and the points they had earned. Training ended after 10 training blocks, each consisting of the 25 training items presented in a random sequence. In the subsequent test phase, participants judged 15 new validation items four times but received no performance feedback.

To motivate participants, they could earn points in every trial. The number of points they earned was a truncated quadratic function of the deviation of their judgment  $j$  from the criterion  $y$ :

$$\text{Points} = 20 - \frac{(j - y)^2}{7.625} \quad (3)$$

At the end of the judgment tasks, the points earned were converted to a monetary bonus (1,500 points = CHF 1). In addition, participants earned a bonus of CHF 3 if they reached 80% of the points in the last training block (corresponding to a root mean square deviation [RMSD] of less than 5.5 in both judgment tasks).

### **Filler Tasks**

The six mostly attention-based filler tasks used during the retention intervals of the memory tests were selected so that they included neither the same type of stimulus material (verbal, numerical, etc.) nor the same items as the memory test. All filler tasks were paper-and-pencil tests. We used the d2 test (Brickenkamp, 2002), the underline “x,” the letter series, the mark numbers divisible by 7, and the number-symbol task from the BIS 4 (Jäger et al., 1997), as well as the letter sets task from the Kit of Factor-Referenced Cognitive Tests (KIT; Ekstrom, French, Harman, & Dermen, 1976). In the d2 test, for instance, participants are asked to cross out all *ds* with two small dashes while ignoring all *ps* or *ds* with more (or fewer) dashes.

### **Procedure**

Participants solved all tasks on one day with a half-hour break between the two sessions. The tasks were presented in the same order to each participant. In the first

session, participants first solved the linear judgment task. Afterward, they solved the operation span, the verb recognition (filler: number-symbol test), the sound identification in noise (filler: letter series), the picture free recall (filler: underline x), and finally completed the symmetry span.

The second session started with the multiplicative judgment task. Afterward, participants completed the reading span, the degraded identification of nouns (filler: mark numbers divisible by 7), the cued number recall task (filler: d2 test), and the speeded presentation of line drawings (filler: letter sets).

## Results

First, we analyzed participants' average performance in the memory and judgment tasks (see Table 2 for descriptive statistics) and modeled participants' judgment strategies. Second, we fitted a measurement model to memory abilities and judgment performance separately. Next, we linked these two measurement models, estimating a structural model that predicts judgment accuracy by memory abilities. Finally, we investigated how strategy execution and strategy selection in the judgment tasks influences the relationship between memory and judgment accuracy.

### Performance Measures

**Performance in the memory tasks.** Performance in the working memory span tasks was comparable to normative data (Redick et al., 2012). Participants achieved a higher partial credit score in the operation and the reading span than in the symmetry span, indicating that they recalled more items in these tasks. In the episodic memory tasks, participants showed a higher recall rate for recognition than for free recall or cued recall. In the implicit memory tasks, participants showed, on average, a higher implicit memory effect in degraded presentation than in identification in noise or speeded presentation. In speeded presentation, participants did not respond faster to the old items at all.



**Performance in the judgment tasks.** First, we assessed how well participants learned to solve the judgment tasks. As an indicator of performance, we calculated the RMSD between participants' judgments in the last training block and the correct criterion, with lower RMSDs indicating higher accuracy. We used Wilcoxon  $z$  tests to compare performance in the judgment tasks, because judgments slightly deviated from normality.

Overall, participants successfully learned to solve the judgment tasks. However, more participants earned a bonus in the multiplicative task (81% of the participants) than in the linear task, 52% of the participants,  $\chi^2(1) = 7.56, p = .006$ . Replicating previous results showing that participants learn to solve multiplicative tasks more accurately than linear tasks (von Helversen & Rieskamp, 2008), participants judged the training items on average more accurately in the multiplicative than in the linear task, Wilcoxon  $z = 4.92, p < .001$ .

Next, we investigated how well people generalized their performance to validation items in the test phase. Performance for validation items was measured as the RSMD between the correct criterion and participants' mean judgment, that is, the judgment for each item averaged over the four presentations. Performance in test was comparable between the linear and the multiplicative task (Wilcoxon  $z = 1.46, p = .145$ ).

### **Modeling of Judgment Strategies**

To investigate which strategy participants relied on, we adopted a cognitive modeling approach. For each participant, we fitted a linear regression model (describing the rule-based strategy), an exemplar model (describing an exemplar-based strategy), and a baseline model (estimating participants' mean judgments) to participants' judgments in the last three training blocks and predicted participants' mean judgments for validation items by using the fitted parameter estimates (von

Helversen & Rieskamp, 2008). This test of predictive fit accounts for model complexity not only in terms of the number of free parameters but also in terms of their functional form (Busemeyer & Wang, 2000).

**Linear model.** In linear regression models, used to mathematically describe rule-based judgment strategies, the importance of each cue is reflected in its cue weight; the higher the cue weights are, the more they influence the judgment. The criterion estimate  $\hat{c}_{p, \text{Linear}}$  of an object  $p$  is the weighted sum of the cue values  $x_{pi}$ :

$$\hat{c}_{p, \text{Linear}} = k + \sum_{i=1}^I w_i \cdot x_{pi} \quad (4)$$

where  $w_i$  are the cue weights for each cue  $i$  and  $k$  is a constant intercept.

**Exemplar model.** To describe the exemplar-based strategy mathematically we used an exemplar model with one free sensitivity parameter (Juslin et al., 2003).<sup>4</sup> In exemplar models, the similarity  $S(p, j)$  between probe  $p$  and exemplar  $j$  is an exponential function of the objects' distance  $d_{pj}$  (Nosofsky & Zaki, 1998):

$$S(p, j) = e^{-d_{pj}}. \quad (5)$$

This distance is determined by summing up the absolute differences between the cue values  $x_{pi}$  of the probe and the cue values  $x_{ji}$  of the exemplar on each cue  $i$  and then weighting this sum by the sensitivity parameter  $h$ .

$$d_{pj} = h \left( \sum_{i=1}^I |x_{pi} - x_{ji}| \right). \quad (6)$$

Correspondingly, the more closely the cue values of the probe and the exemplar match, the smaller the distance is between the objects. The sensitivity parameter expresses how strongly people discriminate among the stored exemplars. A sensitivity parameter close to 0 indicates no discrimination; a high parameter indicates that people specifically remember each exemplar.

The criterion estimate  $\hat{c}_{p,\text{Exemplar}}$  is then determined as the average sum of the similarities weighted by their corresponding criterion values  $c_j$ .

$$\hat{c}_{p,\text{Exemplar}} = \frac{\sum_{j=1}^J S(p,j) \cdot c_j}{\sum_{j=1}^J S(p,j)}, \quad (7)$$

implying that the judgment of a new probe relies upon a similarity-based retrieval of the criterion values associated with each exemplar.

**Model fits.** We measured model fit as the RMSD between model predictions and participants' judgments in the training phase and participants' mean judgments in the test phase, respectively (see Table 3 for fit indices during training and test). A model that perfectly describes participants' judgments would yield an RMSD of 0, whereas a model that, for instance, always overestimates participants' judgments by 9 points would have an RMSD of 9. To compare model fits, we used Wilcoxon signed-rank tests because the RMSDs were not normally distributed.

At the end of training, the baseline model did not provide a good description of participants' judgments in the linear or multiplicative task. The baseline model did worse than the linear and the exemplar model for participants' judgments in the linear task (linear model:  $z = 14.5, p < .001$ ; exemplar model:  $z = 14.3, p < .001$ ) and in the multiplicative task (linear model:  $z = 14.5, p < .001$ ; exemplar model:  $z = 14.2, p < .001$ ). The linear model described participants' judgments overall better than the exemplar model ( $z = 14.5, p < .001$ ) in the linear task, but it did not outperform the exemplar model in the multiplicative task ( $z = 1.5, p = .145$ ).

In the test phase, the linear model also accounted for participants' judgments better than the exemplar model in the linear task ( $z = 11.2, p < .001$ ). In the multiplicative task, the exemplar model predicted participants' judgments slightly more accurately than the linear model ( $z = 4.8, p < .001$ ). Replicating results from

training, the baseline model described participants' judgments worse than the linear model or the exemplar model in the linear task (linear model:  $z = 14.1, p < .001$ ; exemplar model:  $z = 14.2, p < .001$ ) and the multiplicative task (linear model:  $z = 14.0, p < .001$ ; exemplar model:  $z = 14.0, p < .001$ ).

**Strategy classification.** To further examine individual differences in strategy selection, we classified participants as selecting the strategy that led to the smallest RMSD between model predictions and participants' mean judgments. As shown in Figure 1, most participants adapted their strategy to the task. Whereas in the linear task the majority of participants were best described by the linear model ( $n_{\text{Linear}} = 220$ ,  $n_{\text{Exemplar}} = 42$ ,  $n_{\text{Baseline}} = 17$ ), in the multiplicative task most participants were classified as following an exemplar model,  $n_{\text{Linear}} = 99$ ,  $n_{\text{Exemplar}} = 176$ ,  $n_{\text{Baseline}} = 4$ ,  $\chi^2(2) = 136.31, p = .001$ . Indeed, half of the participants (50.2%) were best described by the linear model in the linear task but best described by the exemplar model in the multiplicative task.

To capture how much participants relied on a cue abstraction or an exemplar-based strategy, we also fitted a strategy weight parameter  $w_s$  to participants' judgments in the test phase, excluding participants best described by the baseline model (Hoffmann et al., 2013). Using the optimal parameters for the linear and the exemplar model from the training phase, we estimated the strategy weight by minimizing the RMSD between the models' weighted predictions  $\hat{c}_p$  and participants' mean judgments in the test phase:

$$\hat{c}_p = w_s \cdot \hat{c}_{p,\text{Exemplar}} + (1 - w_s) \cdot \hat{c}_{p,\text{Linear}}. \quad (8)$$

Accordingly, the strategy weight weights the predictions of the exemplar model  $\hat{c}_{p,\text{Exemplar}}$  and the linear model  $\hat{c}_{p,\text{Linear}}$  for the test phase and depends upon only the relative predictive performance of the models. This strategy weight serves as a

measurement tool and can take values between 0 and 1 (see Table 2 for descriptive statistics). A strategy weight above .5 indicates a higher probability for the exemplar model, a strategy weight below .5 a higher probability for the linear model. In the linear task, the strategy weight was on average below .5 (one-sample Wilcoxon test:  $z = 11.3, p < .001$ ), whereas it was larger than .5 in the multiplicative task (one-sample Wilcoxon test:  $z = -3.9, p < .001$ ).

Overall, our results underscore that judgment processes were highly task sensitive (Juslin et al., 2008; Karlsson et al., 2007). Most participants were best described by a rule-based linear judgment strategy in the linear task and by an exemplar-based strategy in the multiplicative task.

### **Structural Equation Modeling**

To understand how judgment performance is grounded in memory abilities, we used structural equation modeling. This approach is particularly strong because it allows testing theories about relations between theoretically well-defined latent constructs extracted from manifest indicators (Hayduk, Cummings, Boadu, Pazderka-Robinson, & Boulianne, 2007). In doing so, structural equation modeling estimates the variance shared among the indicators, correcting the construct for task-specific variance (DeShon, 1998; Tomarken & Waller, 2005). Anderson and Gerbing (1988) recommended first estimating the measurement model that relates the manifest indicators to the latent constructs and then testing the relations between the constructs based on theoretical assumptions.

Model fit is often evaluated based on several fit indices (Iacobucci, 2010; Kline, 2011), among them chi-square ( $\chi^2$ ), the standardized root mean square residual (SRMR), the comparative fit index (CFI), and the root mean square error of approximation (RMSEA). The factor rho coefficient  $\hat{\rho}_{x_i x_i}$  was used to assess the reliability of the constructs (Kline, 2011; Raykov, 2004). To evaluate how the

constructs are related to each other, we compared the hypothesized model to nested competitors by computing  $\chi^2$  difference tests; non-nested models were compared on the Bayesian Information Criterion (BIC). The software package MPLUS (Muthén & Muthén, 2010) was used to estimate the models. Because descriptive data indicated deviations from multivariate normality, we estimated all models using a maximum likelihood estimator with robust standard errors (MLR) and Satorra–Bentler scaled  $\chi^2$  values (scaling factor, SF) for  $\chi^2$  difference tests (Satorra & Bentler, 2001).

**Measurement model for memory abilities.** To establish construct validity, we first estimated a measurement model for memory abilities. We hypothesized that working memory, episodic memory, and implicit perceptual memory constitute three separate constructs, each described by three tests (episodic memory: recognition, free and cued recall; working memory: operation, reading, and symmetry span; implicit perceptual memory: degraded presentation, speeded presentation, and identification in noise). Working memory and episodic memory are typically positively correlated (Brewer & Unsworth, 2012); implicit memory should be uncorrelated with episodic memory (Bruss & Mitchell, 2009) and probably also with working memory.

Table 4 depicts the zero-order correlations between all memory and judgment tasks. Overall, the working memory measures were moderately correlated. Likewise, the episodic memory measures showed small, positive correlations to each other. However, two of three correlations between the implicit memory measures were not different from 0, resulting in an empirically under-identified measurement model for implicit memory. Therefore, we excluded the implicit memory measures from all further analyses, reducing the hypothesized measurement model to two latent factors: working memory and episodic memory.

Indeed, as illustrated in Figure 2, a two-factor latent-variable model assuming that working memory and episodic memory are correlated provided the best fit,  $\chi^2(8)$

= 14.06, SF = 1.00,  $p = .080$ , CFI = .96, RMSEA = .05, SRMR = .04. Omitting the correlation between working memory and episodic memory decreased model fit,  $\chi^2(9) = 21.14$ , SF = 1.06,  $p = .012$ , CFI = .93, RMSEA = .07, SRMR = .07,  $\Delta\chi^2(1) = 5.42$ ,  $p = .020$ . Likewise, assuming only one latent factor for all six memory tasks decreased model fit,  $\chi^2(9) = 37.32$ , SF = 0.95,  $p < .001$ , CFI = .83, RMSEA = .11, SRMR = .06,  $\Delta\chi^2(1) = 38.90$ ,  $p < .001$ . Construct reliabilities in the best fitting model were moderately high for working memory ( $\hat{\rho}_{X_iX_i} = .72$ ) and acceptable for episodic memory ( $\hat{\rho}_{X_iX_i} = .38$ ), considering that episodic memory was assessed with different types of material and tests. Overall, these results replicate one key finding from previous individual difference studies: Working memory and episodic memory are moderately correlated (Brewer & Unsworth, 2012; Del Missier et al., 2013; Unsworth, 2010).

**Measurement model for judgment performance.** The measurement model for judgment performance was particularly interesting because—to our knowledge—whether performance in linear and multiplicative tasks is task specific or depends on a more general ability to learn judgments has not been investigated. To measure performance in both tasks, we used the RMSD between participants' judgments and the correct criterion in each of the four test blocks of the two tasks (see Table 4 for zero-order correlations). Performance in the linear task was assumed to constitute one latent factor and performance in the multiplicative task a second. We then compared three measurement models against each other, assuming that the latent factors are (a) uncorrelated, (b) correlated, or (c) identical; that is, performance over all test blocks in the linear and the multiplicative task can be described by one factor.

As illustrated in Figure 3, a model that assumed a correlation between performances in the linear and multiplicative tasks provided the best fit,  $\chi^2(19) =$

21.87, SF = 1.23,  $p = .291$ , CFI = 1.00, RMSEA = .02, SRMR = .03, suggesting two moderately correlated latent factors. Construct reliability was high for performance in both the linear task ( $\hat{\rho}_{X_i X_i} = .92$ ) and the multiplicative task ( $\hat{\rho}_{X_i X_i} = .90$ ). Omitting the correlation between the latent factors did not harm model fit with regard to CFI (0.99) and RMSEA (.05). However, the other two fit criteria yielded a different picture,  $\chi^2(20) = 33.84$ , SF = 1.24,  $p = .027$ , SRMR = .11,  $\Delta\chi^2(1) = 10.29$ ,  $p = .001$ . A model that assumed a single latent factor was rejected by all fit criteria,  $\chi^2(20) = 571.79$ , SF = 1.15,  $p < .001$ , CFI = 0.53, RMSEA = .31, SRMR = .23.

The small correlation in accuracy between the linear and the multiplicative task yields some evidence that individual differences in judgment performance partly stem from a general ability to solve judgment problems. However, many of the individual differences in accuracy were peculiar to the multiplicative or the linear task, suggesting that distinct processes may account for individual differences in judgment performance.

### **Predicting Judgment Performance With Memory Abilities**

Do individual differences in memory abilities determine how well people solve different judgment tasks? We predicted that participants with higher working memory capacity would make more accurate judgments in the linear task, whereas those with better episodic memory would solve multiplicative tasks more accurately. To test this hypothesis against competing ideas, we combined the measurement model for memory abilities with that for judgment performance into one structural model that assumes a path from working memory to judgment performance in the linear task and a path from episodic memory to judgment performance in the multiplicative task. We compared this model to two alternative models: (1) a null model that assumes memory abilities do not predict judgment performance at all and (2) a full model that



additionally assumes working memory predicts performance in multiplicative tasks and episodic memory predicts performance in linear tasks.

The hypothesized structural model captured the underlying covariance structure very well,  $\chi^2(73) = 67.40$ ,  $SF = 1.04$ ,  $p = .663$ ,  $CFI = 1.00$ ,  $RMSEA = .00$ ,  $SRMR = .04$ , and better than the two alternative models: Assuming no relationship between memory abilities and judgment performance decreased model fit considerably,  $\chi^2(75) = 105.10$ ,  $SF = 1.05$ ,  $p = .012$ ,  $CFI = 0.98$ ,  $RMSEA = .04$ ,  $SRMR = .10$ ,  $\Delta\chi^2(2) = 34.31$ ,  $p < .001$ . Indeed, omitting the path from working memory to judgment performance in the linear task decreased model fit,  $\chi^2(74) = 89.78$ ,  $SF = 1.05$ ,  $p = .102$ ,  $CFI = 0.99$ ,  $RMSEA = .03$ ,  $SRMR = .08$ ,  $\Delta\chi^2(1) = 16.84$ ,  $p < .001$ . Likewise, omitting the path from episodic memory to judgment performance in the multiplicative task decreased model fit,  $\chi^2(74) = 86.89$ ,  $SF = 1.04$ ,  $p = .145$ ,  $CFI = 0.99$ ,  $RMSEA = .03$ ,  $SRMR = .07$ ,  $\Delta\chi^2(1) = 19.50$ ,  $p < .001$ . Finally, the full model that assumed working memory and episodic memory are both important for predicting judgment performance in the linear and the multiplicative task did not outperform the hypothesized model,  $\chi^2(71) = 63.76$ ,  $SF = 1.04$ ,  $p = .72$ ,  $CFI = 1.00$ ,  $RMSEA = .00$ ,  $SRMR = .03$ ,  $\Delta\chi^2(2) = 3.52$ ,  $p = .172$ .

In line with our hypothesis, the resulting structural model (Figure 4) shows that people with higher working memory capacity solved linear tasks more accurately than people with lower working memory capacity, whereas people with better episodic memory solved multiplicative tasks better than those with bad episodic memory.<sup>5</sup> We next investigated if memory abilities also influence strategy selection.

### **Tracing the Path From Memory Abilities to Judgment Performance Through Judgment Strategies**

**Strategy selection.** In the Introduction we outlined that memory abilities may change strategy selection in two ways. Working memory may make it more likely that people quickly detect the task-appropriate judgment strategy; accordingly, working memory should predict strategy selection in the linear and the multiplicative task, and strategy selection, in turn, should predict judgment accuracy. Yet it is possible that people with better episodic memory rely more on their capabilities and select exemplar-based strategies more often in the multiplicative task. Accordingly, episodic memory should predict strategy selection in the multiplicative task and, in turn, judgment accuracy.

To investigate how memory abilities affect strategy selection and judgment accuracy, we relied on mediation analyses. If memory abilities influence judgment accuracy by altering the strategy, then strategy selection should mediate the relationship between memory abilities and judgment performance. We compared a null model that assumed strategy selection does not mediate the relationship between memory abilities and judgment accuracy against different mediator models. Alternative models proposed that strategy selection mediates the relationship (a) between episodic memory and performance only in the multiplicative task, (b) between working memory and performance in the linear task, or (c) between working memory and performance in the multiplicative task.

To conduct these analyses, we used the continuous strategy weight  $w_s$  as the mediator (Equation 8). Because the strategy weight indicates only how much participants relied on an exemplar-based strategy or a rule-based strategy, participants classified as following a baseline model in the linear or the multiplicative task were coded as missing on that variable. To avoid excluding all their data, we used a full information maximum likelihood approach to estimate the structural model (Tomarken & Waller, 2005).

Overall, the best fitting structural model assumed that episodic memory predicts strategy selection in the multiplicative task and this choice, in turn, influences judgment accuracy in the multiplicative task,  $\chi^2(100) = 94.94$ , SF = 1.03,  $p = .624$ , CFI = 1.00, RMSEA = .00, SRMR = .05, BIC = 13,928. This model fit significantly better than a model that did not assume a path from memory abilities to strategy selection or from strategy selection to judgment performance,  $\chi^2(102) = 186.60$ , SF = 1.04,  $p < .001$ , CFI = .95, RMSEA = .06, SRMR = .10, BIC = 14,012,  $\Delta\chi^2(2) = 83.75$ ,  $p < .001$ . Model fit was also improved by assuming that working memory predicts strategy selection in the multiplicative task,  $\chi^2(100) = 99.45$ , SF = 1.03,  $p = .497$ , CFI = 1.00, RMSEA = .00, SRMR = .05, BIC = 13,933,  $\Delta\chi^2(2) = 79.64$ ,  $p < .001$ . However, this model yields a higher BIC than the model that predicts strategy selection in the multiplicative task with episodic memory,  $\Delta\text{BIC} = 5$ , suggesting that episodic memory predicts strategy selection in the multiplicative task slightly better than working memory. Considering both predictors simultaneously also did not increase model fit,  $\chi^2(99) = 93.78$ , SF = 1.03,  $p = .629$ , CFI = 1.00, RMSEA = .00, SRMR = .04,  $\Delta\chi^2(1) = 1.17$ ,  $p = .280$ . Model fit was not improved by assuming that strategy selection mediates the relationship between working memory and judgment accuracy in the linear task,  $\chi^2(100) = 182.30$ , SF = 1.03,  $p < .001$ , CFI = 0.95, RMSEA = .05, SRMR = .09, BIC = 14,019,  $\Delta\chi^2(2) = 4.27$ ,  $p = .118$ .<sup>6</sup>

As illustrated in Figure 5, the best fitting structural model shows that strategy selection partly mediated the relationship between episodic memory and judgment performance in the multiplicative task. People with better episodic memory were more likely to select an exemplar-based strategy in the multiplicative task, and this strategy change increased judgment accuracy in the multiplicative task ( $r = -.16$  for the indirect effect,  $p < .001$ ). Better episodic memory still predicted higher judgment accuracy, but

the standardized regression weight dropped from  $r = -.43$  to  $r = -.27$  when the strategy weight in the multiplicative task (called “strategy” in the structural model) was added. In contrast, higher working memory capacity did not increase the probability of selecting a rule-based strategy in the linear task and strategy selection did not affect judgment performance in the linear task.

**Strategy execution.** In the Introduction we argued that memory abilities may predict judgment performance because memory abilities improve strategy execution. Specifically, high working memory capacity may help people execute rule-based strategies, and in turn, strategy execution may mediate the relationship between working memory capacity and judgment accuracy in the linear task. In contrast, episodic memory may help people execute exemplar-based strategies, and in turn, strategy execution may mediate the relationship between episodic memory and judgment accuracy in the multiplicative task. To further test these hypotheses, we examined how strategy execution contributes to the relationship between memory and judgment. As an indicator for strategy execution in the linear and the multiplicative task, we used the model fit resulting from the estimation of the strategy weight (Equation 8) that is the minimal RMSD between the weighted model predictions  $\hat{c}_p$  and participants’ mean judgments in the test phase. If the strategy weight is 0, the strategy execution measure equates to the predictive fit of the linear model; if the strategy weight is 0.5, it reflects the combined fit of both models. Consequently, the strategy execution measure determines how consistently people transfer the strategy learned in training to validation items in test given the strategy weight.

To understand how strategy execution is related to memory and judgment accuracy, we again conducted mediation analyses. Matching the analysis for strategy selection, we estimated a null model that assumed strategy execution does not mediate the relationship between memory abilities and judgment accuracy. We compared this

model to different competitors that assumed strategy execution mediates the relationship between (a) working memory and performance only in the linear task, or (b) episodic memory and performance only in the multiplicative task.

In the best fitting structural model, strategy execution mediated the relationship between working memory and judgment accuracy in the linear task,  $\chi^2(100) = 102.57$ , SF = 1.05,  $p = .410$ , CFI = 1.00, RMSEA = .01, SRMR = .04, BIC = 15,249.

According to this model, working memory predicts strategy execution in the linear task; hence, the more closely participants followed the strategy learned in training, the more accurate were their judgments. Comparing this model to the null model and thus discarding the indirect effect of strategy execution in the linear task significantly harmed the fit of the structural model,  $\chi^2(102) = 208.13$ , SF = 1.05,  $p < .001$ , CFI = .94, RMSEA = .06, SRMR = .10,  $\Delta\chi^2(2) = 105.56$ ,  $p < .001$ , BIC = 15,349. A structural model assuming that strategy execution mediates the relationship between episodic memory and accuracy in the multiplicative task did not improve model fit compared to the null model,  $\chi^2(100) = 202.69$ , SF = 1.05,  $p < .001$ , CFI = 0.94, RMSEA = .06, SRMR = .10,  $\Delta\chi^2(2) = 5.26$ ,  $p = .072$ , BIC = 15,356. Likewise, this model yields a higher BIC than the model predicting strategy execution with working memory in the linear task,  $\Delta\text{BIC} = 107$ .<sup>7</sup>

Figure 6 shows the resulting structural model. In this model, working memory capacity again directly predicts judgment accuracy in the linear task, but to a smaller extent (the standardized regression weight fell from  $r = -.35$  to  $r = -.24$ ). Strategy execution mediates this relationship between working memory and judgment accuracy. Higher working memory capacity facilitates executing the learned strategy in linear tasks, and strategy execution, in turn, predicts how accurately people make judgments in linear tasks ( $r = -.11$  for the indirect effect,  $p = .019$ ). In the

multiplicative task, however, episodic memory does not predict how well people execute a learned strategy, and strategy execution does not lead to more accurate judgments.

### **General Discussion**

Working memory and long-term memory are indispensable for many everyday activities. In fact, working memory capacity predicts performance in a wide range of cognitive tasks ranging from reading (Daneman & Carpenter, 1980) to reasoning (Kane et al., 2004; Kyllonen & Christal, 1990) and also predicts everyday cognitive failures (Unsworth et al., 2012). Likewise, episodic memory has proven useful as an indicator of general intelligence (Jäger et al., 1997). However, little attention has been paid to how various memory abilities influence judgment and decision making (Ashby & O'Brien, 2005; Del Missier et al., 2013; Tomlinson et al., 2011). We sought to fill this gap by investigating how working memory and episodic memory promote judgment strategies and performance in two judgment tasks: a linear task that can best be solved by a rule-based cue abstraction strategy and a multiplicative task in which people should rely more often on an exemplar-based strategy. As predicted, we found that working memory capacity was linked to judgment accuracy in linear tasks in which most people tried to follow abstract rules. In contrast, episodic memory was related to judgment accuracy in multiplicative tasks in which most people relied on exemplar-based strategies. Furthermore, working memory did not predict performance in multiplicative tasks above episodic memory. Thus, largely in line with theories in judgment and decision making—and even more with categorization theories (Ashby & O'Brien, 2005; Juslin et al., 2008; Smith et al., 1998)—these results suggest that rule-based and exemplar-based strategies tap into different memory abilities.

The idea that memory abilities may affect not only how well people execute a strategy but also what strategy they choose has attracted much attention (Bröder, 2003;

Mata, von Helversen et al., 2012). Past research particularly focused on the role of working memory for adaptive strategy selection (Craig & Lewandowsky, 2012; Lewandowsky et al., 2012; McDaniel, Cahill, Robbins, & Wiener, 2014). In our study, we found that episodic memory predicted the probability with which a person was best described by an exemplar-based strategy in the multiplicative task. Working memory capacity, however, did not affect strategy selection above episodic memory, suggesting that working memory could be less important for adaptive strategy selection than previously assumed (Lewandowsky et al., 2012; Mata, Pachur et al., 2012). Next, we discuss in detail how memory abilities may influence judgment strategies and performance.

### **The Influence of Memory Abilities on Rule-based Strategies**

Rule-based strategies have often been understood as serial, capacity-constrained hypothesis-testing processes that demand high working memory capacity (Ashby & O'Brien, 2005; Brehmer, 1994; Juslin et al., 2003, 2008). Supporting the idea that working memory capacity is indispensable for making rule-based judgments, we found that working memory was related to judgment accuracy in linear tasks in which participants' judgments were, overall, best described by a linear rule. This result resonates well with previous findings showing that successfully adopting a rule-based strategy is impeded by cognitive load (Filoteo et al., 2010; Hoffmann et al., 2013). Theoretically, two major components of rule-based strategies contribute to the relationship between working memory capacity and judgment accuracy. First, abstracting linear rules may require maintaining the previous judgment object in working memory and comparing it to the current judgment object (Juslin et al., 2008; Pachur & Olsson, 2012). Second, executing a rule-based strategy may involve mental updating of the judgment estimate and inhibiting irrelevant cue information. In line with the latter idea, we found that working memory capacity promoted executing the

chosen strategy more consistently in linear tasks, and strategy execution, in turn, predicted judgment accuracy. This finding matches previous research suggesting that working memory capacity influences how accurately people apply decision rules (Del Missier et al., 2013).

Our results, however, seem to contradict findings by Rolison, Evans, Walsh, and Dennis (2011) who suggested that working memory capacity is required only for learning negative, and not positive relationships between the cues and the criterion. In contrast, we found that working memory also predicted how successful people were at learning positive cue–criterion relationships. However, our study also used more predictive cues than Rolison et al.’s study (four cues instead of two). One explanation could be that both negative cue–criterion relationships and a higher number of cues make testing of alternative hypotheses more difficult. Possibly, people with low working memory capacity can still test hypotheses about two cues, whereas only high working memory capacity allows people to consider more alternative hypotheses (Dougherty & Hunter, 2003).

Episodic memory, in our study, did not directly predict judgment accuracy in linear tasks, suggesting that episodic memory is less important than working memory capacity for making rule-based judgments. However, memory skills are not independent of each other. Replicating findings from memory research (Del Missier et al., 2013; Mogle et al., 2008; Unsworth, 2010), we found that working memory and episodic memory are moderately correlated, probably reflecting that working memory is needed to encode and retrieve information from long-term memory. Consequently, episodic memory was indirectly related to accuracy in linear tasks through its correlation with working memory ( $r = -.14$ , with the correlation computed as the product of the correlation between working memory and episodic memory,  $r = .41$ , and the standardized coefficient from working memory to linear judgment accuracy,



standardized coefficient =  $-.35$ ). Possibly, this indirect relationship suggests that episodic memory is still needed to retrieve cue weights when making a judgment.

### **The Influence of Memory Abilities on Exemplar-based Strategies**

Surprisingly few studies have investigated how episodic memory is linked to strategies and performance in judgments or decision making. Our study emphasizes how important episodic memory is for making exemplar-based judgments. We found clear evidence that episodic memory predicts judgment accuracy in multiplicative tasks in which participants' judgments were best described by an exemplar-based strategy. This result is in line with previous studies suggesting that people engage in strategic memorization when adopting exemplar-based strategies (Juslin et al., 2008; Olsson et al., 2006) and further supports the theoretical link between episodic memory trace and exemplar models (Hintzman, 1984, 1986).

In contrast to the linear task, we found no direct link between working memory capacity and judgments in the multiplicative task suggesting that working memory does not contribute to performance above episodic memory. Thus, even if exemplar processes rely on controlled retrieval (e.g., Karlsson et al., 2008) they appear to require less working memory than a rule-based strategy. This result may contradict previous findings that working memory helps solve different judgment and categorization tasks (Craig & Lewandowsky, 2012; Lewandowsky, 2011; Lewandowsky et al., 2012; Weaver & Stewart, 2012).

Our results might differ from the previous literature because we investigated how successfully people generalized their performance to new items, whereas previous studies focused mostly on the learning process. In Lewandowsky's (2011) study, for instance, a learning parameter best captured variations in working memory capacity across six different categorization tasks. In addition, we assessed judgment performance—because of time restrictions—with only two tasks, using accuracy in the

four test blocks as manifest indicators. Accordingly, our measurement focused more strongly on variance specific to each judgment task, whereas past research concentrated on the variance shared among different judgment or categorization tasks (Craig & Lewandowsky, 2012; Lewandowsky, 2011; Lewandowsky et al., 2012; Weaver & Stewart, 2012). Hence, it is possible that learning in rule- and exemplar-based judgment tasks requires working-memory capacity, whereas executing a learned strategy depends on working memory capacity only for rule-based judgments. However, as mentioned above, working memory capacity was moderately correlated with episodic memory in our study. Accordingly, working memory was helpful for solving not only linear tasks but also multiplicative tasks: Higher working memory capacity predicted higher judgment accuracy in the multiplicative task through its connection to episodic memory ( $r = -.17$ ). In sum, successfully solving judgment tasks relies on the interplay between episodic memory and working memory—an interpretation that is generally in line with the idea that learning in a huge variety of judgment tasks depends on working memory capacity (Weaver & Stewart, 2012).

### **Memory Abilities and Strategy Use**

In the past decade, research has focused mostly on task characteristics as a determinant of judgment strategies (Juslin et al., 2003, 2008; Karlsson et al., 2007; von Helversen et al., 2013; von Helversen & Rieskamp, 2009). Consistent with prior research, we found that most participants relied on a rule-based strategy in a linear task and shifted to exemplar-based strategies in multiplicative tasks (Hoffmann et al., 2013; Juslin et al., 2008; Karlsson et al., 2007). However, individual differences, such as age or intelligence, can also drive shifts between different types of strategies (Bröder, 2003; Mata, von Helversen et al., 2012). Specifically, we argued that memory abilities may influence not only how well people execute a strategy but also which strategies they select (Beach & Mitchell, 1978; Lemaire & Siegler, 1995; Mata,

Pachur, et al., 2012). Whereas neither working memory capacity nor episodic memory influenced strategy selection in the linear task, episodic memory fostered the probability of selecting an exemplar strategy in the multiplicative task. Furthermore, strategy selection partly mediated the relationship between episodic memory and judgment performance. This result dovetails with the idea that memory abilities may reduce the costs associated with a strategy and, in turn, increase the preference for employing a specific strategy (Beach & Mitchell, 1978; Payne et al., 1993; Rieskamp & Otto, 2006).

Following the strategy selection approach, however, working memory capacity should also have predicted the extent to which people select a rule-based strategy in the linear task. One reason we did not find this relationship could be that rule-based strategies act as a default (Karlsson et al., 2008; Olsson et al., 2006). In line with this argumentation, few people chose an exemplar strategy in the linear tasks. Consequently, only engaging in exemplar-based memorization would require an active choice, whereas the success of rule-based strategies may depend more on the effort needed to execute the strategy. This explanation is supported by the finding that working memory capacity predicted how well the learned strategy was executed in the linear task, suggesting that the inability to accurately use a strategy does not necessarily lead to a strategy shift. In contrast, how well the learned strategy was executed in the multiplicative task was unrelated to episodic memory, suggesting that those participants who did not shift to the task-appropriate exemplar-based strategy nevertheless applied the rules they learned consistently.

In sum, our results demonstrate that episodic memory plays an important role in strategy selection (Mata, von Helversen et al., 2012) but do not provide any evidence that working memory capacity—as previously suggested—predicts more adaptive strategy selection (Bröder, 2003; Lewandowsky et al., 2012; Mata, Pachur, et

al., 2012). These results emphasize that reducing strategy selection to a question of working memory capacity probably oversimplifies the idea of adaptive strategy use.

### **Future Directions**

We established a link between explicit memory and judgment strategies. One puzzle that remains to be solved is the degree to which judgments rely upon implicit memory. Specifically, scholars have heatedly debated if implicit perceptual memory supports learning in similarity-based categorizations (Knowlton & Squire, 1993; Nosofsky & Zaki, 1998; Palmeri & Flanery, 1999; Smith & Grossman, 2008). In our study, we used several established tasks to measure implicit perceptual memory and to examine its relation to exemplar-based judgments. However, correlations between implicit memory tasks were low so we did not include implicit memory in the analysis. Accordingly, the relation between implicit perceptual memory and exemplar-based judgments remains unclear.

A related unresolved debate deals with the question of how implicit procedural memory contributes to learning in judgment and categorization (Ashby & Maddox, 2005; Newell et al., 2011). Procedural memory underlies the learning of motor skills (Squire & Zola, 1996; Willingham, 1998) and may also contribute to learning “structured categories containing many exemplars that could not be easily learned via a logical reasoning process” (Ashby & O’Brien, 2005, p. 86). In these information-integration tasks, the optimal strategy is difficult to verbalize and learning requires many repetitions (Ashby & Maddox, 2005). In line with the idea that procedural learning underlies information integration, it has been found that disrupting motor processing harms performance in information-integration tasks more strongly than performance with rule-based categorizations (Ashby, Ell, & Waldron, 2003; Maddox, Bohil, & Ing, 2004; but see Zaki & Kleinschmidt, 2014).

Structurally, information-integration tasks in categorization are most similar to linear judgment tasks, which means implicit procedural memory might help learning in these tasks. Judgment research has instead suggested that people test specific hypotheses when learning to solve linear tasks (Brehmer, 1994; Juslin et al., 2008). Furthermore, it has been shown that participants acquire explicit knowledge about the cues' importance when solving information-integration tasks (Lagnado et al., 2006). These results complement our finding that performance in the linear task was predicted by working memory, indicating that explicit reasoning was involved (Ashby & O'Brien, 2005). The degree to which implicit procedural memory plays a role in learning to solve judgment tasks is an open question that should be tackled by future research.

Another unresolved problem is the question of how far the relationship between memory abilities and performance found in test can be generalized to the learning phase. Performance at test strongly depends on how well the task was learned, suggesting that those memory abilities that influence test performance should also matter during learning. Yet how people learn to rely on linear rules or exemplars is barely understood and only a few attempts have been made to understand and mathematically describe the learning process (e.g., Kelley & Busemeyer, 2008; Lagnado et al., 2006; Speekenbrink & Shanks, 2010).

A prominent model that is thought to capture cue learning on a trial-by-trial basis is the Gluck and Bower (1988) model (see also Kelley & Busemeyer, 2008), which assumes that all cue weights are updated with the same learning speed—a learning process possibly supported by implicit procedural learning. Thus, learning to make judgments might be driven largely by implicit procedural memory and be less dependent on working memory. In contrast, a capacity-limited hypothesis-testing process may consider and update only one hypothesis in a trial. In line with this

hypothesis, Markant and Gureckis (2014) found in rule-based categorizations that participants with the opportunity to sample information according to their current hypothesis outperformed those who saw the same stimuli but could not actively choose them. Thus, a learning process that relies on hypothesis testing may depend even more strongly on working memory than applying a rule-based strategy during test.

Similarly, although we found that performance in a multiplicative task depends more on episodic memory and less on working memory, it is possible that learning to solve a multiplicative task would more strongly involve working memory.

Successfully learning to rely on exemplar memory possibly requires not only the controlled retrieval of exemplars, but also strengthening the association between the exemplar and the outcome criteria, a process that could benefit from working memory capacity (Lewandowsky, 2011). In sum, although our study offers some insights into the relationship between memory abilities and judgment performance, it is far less clear what abilities are involved in learning these tasks.

## **Conclusions**

Twenty years ago, Elke Weber and colleagues (1995) reminded us that we should not forget memory when thinking about how people make judgments. Our results suggest that different judgment strategies take advantage of specific memory processes: Whereas rule-based strategies draw on working memory capacity, exemplar-based strategies exploit encoding and retrieval processes in episodic memory. Thus, knowledge about working memory and long-term memory may help explain how people successfully solve judgment tasks that range from daily judgments such as estimating the probability of rainfall to professional judgments such as judging the quality of a job candidate.

## Footnotes

1. In a pilot study, 12 participants rated 100 German sentences for plausibility. Only highly plausible or implausible sentences were included in the final reading span.
2. In a pilot study, we included a threshold procedure using 40 independent nouns. The results showed that participants correctly identified half of the nouns using a mask with nine black squares so that 56% of the noun was masked.
3. To assure that old and new sounds were equally easy to identify among distractors, we conducted a pilot study with 24 subjects: Half of the participants heard half of the sounds without noise in the study phase; the other half heard the remaining sounds in the study phase. Afterward, old and new sounds were presented in noise and participants were asked to identify them among two distractors. For the final experiment, old and new sounds were matched on performance for old sounds.
4. We also fitted an exemplar model with four attention parameters to participants' judgments. However, replicating results from previous studies (Hoffmann et al., 2013; von Helversen & Rieskamp, 2008), this model failed to outperform an exemplar model with one parameter in predicting participants' judgments for validation items in either the linear task (RMSD = 5.3) or the multiplicative task (RMSD = 5.9).
5. Judgment accuracy was measured in RMSD with lower RMSD indicating more accurate judgments. Accordingly, negative correlations imply that higher working memory predicts higher judgment accuracy.
6. A structural model testing for all mediation effects simultaneously led to similar conclusions: Overall, this model achieved a good fit,  $\chi^2(97) = 89.63$ , SF = 1.03,  $p = .690$ , CFI = 1.00, RMSEA = .00, SRMR = .04, BIC = 13,940. Testing for mediation effects suggested only an indirect effect of episodic memory over strategy selection in the multiplicative task on judgment accuracy in this task ( $r = -.13$ ,  $p =$

.013). Neither the indirect effect of working memory on judgment accuracy in the multiplicative task nor the indirect effect of working memory on judgment accuracy in the linear task was significant (multiplicative:  $r = -.05$ ,  $p = .282$ ; linear:  $r = -.003$ ,  $p = .743$ ).

7. Testing for both mediation effects simultaneously led to similar conclusions:

Overall, this model achieved a good fit,  $\chi^2(98) = 96.75$ ,  $SF = 1.05$ ,  $p = .517$ ,  $CFI = 1.00$ ,  $RMSEA = .00$ ,  $SRMR = .04$ ,  $BIC = 15,253$ . Testing for mediation effects suggested only an indirect effect of working memory on judgment accuracy in the linear task through strategy execution ( $r = -.11$ ,  $p = .021$ ). The indirect effect of episodic memory on judgment accuracy in the multiplicative task was not significant ( $r = -.02$ ,  $p = .389$ ).



## References

- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103, 411–423. doi:10.1037/0033-2909.103.3.411
- Ashby, F. G., Ell, S. W., & Waldron, E. M. (2003). Procedural learning in perceptual categorization. *Memory & Cognition*, 31, 1114–1125. doi:10.3758/BF03196132
- Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology*, 56, 149–178. doi:10.1146/annurev.psych.56.091103.070217
- Ashby, F. G., & O'Brien, J. B. (2005). Category learning and multiple memory systems. *Trends in Cognitive Sciences*, 9, 83–89. doi:10.1016/j.tics.2004.12.003
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, pp. 89–195). New York, NY: Academic Press.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning & motivation* (Vol. 8, pp. 47–89). New York, NY: Academic Press.
- Baddeley, A., Lewis, V., Eldridge, M., & Thomson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General*, 113, 518–540. doi:10.1037//0096-3445.113.4.518
- Beach, L. R., & Mitchell, T. R. (1978). A contingency model for the selection of decision strategies. *The Academy of Management Review*, 3, 439–449.

- Berry, C. J., Shanks, D. R., Speekenbrink, M., & Henson, R. N. A. (2012). Models of recognition, repetition priming, and fluency: Exploring a new framework. *Psychological Review*, 119, 40–79. doi:10.1037/a0025464
- Brehmer, B. (1994). The psychology of linear judgement models. *Acta Psychologica*, 87, 137–154. doi:10.1016/0001-6918(94)90048-5
- Brewer, G. A., & Unsworth, N. (2012). Individual differences in the effects of retrieval from long-term memory. *Journal of Memory and Language*, 66, 407–415. doi:10.1016/j.jml.2011.12.009
- Brickenkamp, R. (2002). *Test d2. Aufmerksamkeits-Belastungs-Test*. Göttingen, Germany: Hogrefe.
- Bröder, A. (2003). Decision making with the “adaptive toolbox”: Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 611–625. doi:10.1037/0278-7393.29.4.611
- Brown, S., & Craik, F. (2000). Encoding and retrieval of information. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 93–107). New York, NY: Oxford University Press.
- Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York, NY: Wiley.
- Brunswik, E. (1956). *Perception and the representative design of psychological experiments* (2nd ed.). Berkeley, CA: University of California Press.
- Bruss, P. J., & Mitchell, D. B. (2009). Memory systems, processes, and tasks: Taxonomic clarification via factor analysis. *The American Journal of Psychology*, 122, 175–189.

- Buchner, A., & Brandt, M. (2003). Further evidence for systematic reliability differences between explicit and implicit memory tests. *The Quarterly Journal of Experimental Psychology A*, 56, 193–209. doi:10.1080/02724980244000260
- Buchner, A., & Wippich, W. (2000). On the reliability of implicit and explicit memory measures. *Cognitive Psychology*, 40, 227–259. doi:10.1006/cogp.1999.0731
- Busemeyer, J. R., & Wang, Y.-M. (2000). Model comparisons and model selections based on generalization criterion methodology. *Journal of Mathematical Psychology*, 44, 171–189. doi:10.1006/jmps.1999.1282
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132, 354–380. doi:10.1037/0033-2909.132.3.354
- Chiu, C. Y. (2000). Specificity of auditory implicit and explicit memory: Is perceptual priming for environmental sounds exemplar specific? *Memory & Cognition*, 28, 1126–1139. doi:10.3758/BF03211814
- Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12, 769–786. doi:10.3758/BF03196772
- Cooksey, R. W. (1996). The methodology of social judgement theory. *Thinking & Reasoning*, 2, 141–174. doi:10.1080/135467896394483
- Craig, S., & Lewandowsky, S. (2012). Whichever way you choose to categorize, working memory helps you learn. *The Quarterly Journal of Experimental Psychology*, 65, 439–464. doi:10.1080/17470218.2011.608854
- Craik, F. I., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human

- memory. *Journal of Experimental Psychology: General*, 125, 159–180.  
doi:10.1037/0096-3445.125.2.159
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450–466. doi:10.1016/S0022-5371(80)90312-6d
- Davis, T., Love, B. C., & Preston, A. R. (2012). Learning the exception to the rule: Model-based fMRI reveals specialized representations for surprising category members. *Cerebral Cortex*, 22, 260–273. doi:10.1093/cercor/bhr036
- Del Missier, F., Mäntylä, T., & Bruine de Bruin, W. (2010). Executive functions in decision making: An individual differences approach. *Thinking & Reasoning*, 16, 69–97. doi:10.1080/13546781003630117
- Del Missier, F., Mäntylä, T., & Bruine de Bruin, W. (2012). Decision-making competence, executive functioning, and general cognitive abilities. *Journal of Behavioral Decision Making*, 25, 331–351. doi:10.1002/bdm.731
- Del Missier, F., Mäntylä, T., Hansson, P., Bruine de Bruin, W., Parker, A. M., & Nilsson, L.-G. (2013). The multifold relationship between memory and decision making: An individual-differences study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 1344–1364.  
doi:10.1037/a0032379
- DeShon, R. (1998). A cautionary note on measurement error corrections in structural equation models. *Psychological Methods*, 3, 412–423.
- Dew, I. T. Z., & Cabeza, R. (2011). The porous boundaries between explicit and implicit memory: Behavioral and neural evidence. *Annals of the New York Academy of Sciences*, 1224, 174–190. doi:10.1111/j.1749-6632.2010.05946.x

- Dougherty, M. R. P., Gettys, C. F., & Ogden, E. E. (1999). MINERVA-DM: A memory processes model for judgments of likelihood. *Psychological Review*, 106, 180–209.
- Dougherty, M. R. P., & Hunter, J. E. (2003). Hypothesis generation, probability judgment, and individual differences in working memory capacity. *Acta Psychologica*, 113, 263–282. doi:10.1016/S0001-6918(03)00033-7
- Einhorn, H. J., Kleinmuntz, D. N., & Kleinmuntz, B. (1979). Linear regression and process-tracing models of judgment. *Psychological Review*, 86, 465–485. doi:10.1037/0033-295X.86.5.465
- Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.
- Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, 127, 107–140. doi:10.1037/0096-3445.127.2.107
- Estes, W. K. (1986). Memory storage and retrieval processes in category learning. *Journal of Experimental Psychology: General*, 115, 155–174.
- Filoteo, J. V., Lauritzen, J. S., & Maddox, W. T. (2010). Removing the frontal lobes: The effects of engaging executive functions on perceptual category learning. *Psychological Science*, 21, 415–423. doi:10.1177/0956797610362646
- Fleischman, D. A., Wilson, R. S., Gabrieli, J. D. E., Bienias, J. L., & Bennett, D. A. (2004). A longitudinal study of implicit and explicit memory in old persons. *Psychology and Aging*, 19, 617–625. doi:10.1037/0882-7974.19.4.617
- Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple heuristics that make us smart*. New York, NY: Oxford University Press.

- Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, 117, 227–247. doi:10.1037/0096-3445.117.3.227
- Graves, L., & Karren, R. (1992). Interviewer decision processes and effectiveness: An experimental policy-capturing investigation. *Personnel Psychology*, 45, 313–340. doi:10.1111/j.1744-6570.1992.tb00852.x
- Gygi, B., & Shafiro, V. (2010). Development of the database for environmental sound research and application (DESRA): Design, functionality, and retrieval considerations. *EURASIP Journal on Audio, Speech, and Music Processing*, 2010, 1–12. doi:10.1155/2010/654914
- Hager, W., & Hasselhorn, M. (Eds.). (1994). *Handbuch deutschsprachiger Wortnormen*. Göttingen: Hogrefe.
- Hahn, U., & Chater, N. (1998). Similarity and rules: Distinct? Exhaustive? Empirically distinguishable? *Cognition*, 65, 197–230. doi:10.1016/S0010-0277(97)00044-9
- Hayduk, L., Cummings, G., Boadu, K., Pazderka-Robinson, H., & Boulianne, S. (2007). Testing! Testing! One, two, three—Testing the theory in structural equation models! *Personality and Individual Differences*, 42, 841–850. doi:10.1016/j.paid.2006.10.001
- Helsdingen, A. S., van Gog, T., & van Merriënboer, J. J. G. (2011). The effects of practice schedule on learning a complex judgment task. *Learning and Instruction*, 21, 126–136. doi:10.1016/j.learninstruc.2009.12.001
- Hintzman, D. L. (1984). MINERVA 2 : A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers*, 16, 96–101.
- Hintzman, D. L. (1986). “Schema abstraction” in a multiple-trace memory model. *Psychological Review*, 93, 411–428. doi:10.1037//0033-295X.93.4.411

- Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *95*, 528–551.  
doi:10.1037//0033-295X.95.4.528
- Hoffmann, J. A., von Helversen, B., & Rieskamp, J. (2013). Deliberation's blindsight: How cognitive load can improve judgments. *Psychological Science*, *26*, 869–879. doi:10.1177/0956797612463581
- Hoffmann, J. A., von Helversen, B., & Rieskamp, J. (2014). *Similar task features shape judgment and categorization processes*. Manuscript submitted for publication.
- Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology*, *24*, 1–55.
- Iacobucci, D. (2010). Structural equations modeling: Fit indices, sample size, and advanced topics. *Journal of Consumer Psychology*, *20*, 90–98.  
doi:10.1016/j.jcps.2009.09.003
- Jäger, A. O., Süß, H.-M., & Beauducel, A. (1997). *Berliner Intelligenzstruktur-Test. Form 4*. Göttingen, Germany: Hogrefe.
- Juslin, P., Karlsson, L., & Olsson, H. (2008). Information integration in multiple cue judgment: A division of labor hypothesis. *Cognition*, *106*, 259–298.  
doi:10.1016/j.cognition.2007.02.003
- Juslin, P., Olsson, H., & Olsson, A.-C. (2003). Exemplar effects in categorization and multiple-cue judgment. *Journal of Experimental Psychology: General*, *132*, 133–156. doi:10.1037/0096-3445.132.1.133
- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning.

*Journal of Experimental Psychology: General*, 133, 189–217.

doi:10.1037/0096-3445.133.2.189

Karlsson, L., Juslin, P., & Olsson, H. (2007). Adaptive changes between cue abstraction and exemplar memory in a multiple-cue judgment task with continuous cues. *Psychonomic Bulletin & Review*, 14, 1140–1146.

Karlsson, L., Juslin, P., & Olsson, H. (2008). Exemplar-based inference in multi-attribute decision making: Contingent, not automatic, strategy shifts? *Judgment and Decision Making*, 3, 244–260.

Kelley, H., & Busemeyer, J. (2008). A comparison of models for learning how to dynamically integrate multiple cues in order to forecast continuous criteria. *Journal of Mathematical Psychology*, 52, 218–240.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York, NY: Guilford.

Knowlton, B. J. (1999). What can neuropsychology tell us about category learning? *Trends in Cognitive Sciences*, 3, 123–124. doi:10.1016/S1364-6613(99)01292-9

Knowlton, B. J., & Squire, L. R. (1993). The learning of categories: Parallel brain systems for item memory and category knowledge. *Science*, 262, 1747–1749. doi:10.1126/science.8259522

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, 14, 389–433. doi:10.1016/S0160-2896(05)80012-1

Lagnado, D. A., Newell, B., Kahan, S., & Shanks, D. R. (2006). Insight and strategy in multiple cue learning. *Journal of Experimental Psychology: General*, 135, 162–183. doi:10.1037/0096-3445.135.2.162



- Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change: Contributions to children's learning of multiplication. *Journal of Experimental Psychology: General*, 124, 83–97. doi:10.1037//0096-3445.124.1.83
- Lewandowsky, S. (2011). Working memory capacity and categorization: Individual differences and modeling. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 720–738. doi:10.1037/a0022639
- Lewandowsky, S., Yang, L. X., Newell, B. R., & Kalish, M. L. (2012). Working memory does not dissociate between different perceptual categorization tasks. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 38, 881–904. doi:10.1037/a0027298
- Maddox, W. T., Bohil, C. J., & Ing, A. D. (2004). Evidence for a procedural- learning-based system in perceptual category learning. *Psychonomic Bulletin & Review*, 11, 945–952. doi:10.3758/BF03196726
- Marewski, J. N., & Schooler, L. J. (2011). Cognitive niches: An ecological model of strategy selection. *Psychological Review*, 118, 393–437. doi:10.1037/a0024143
- Markant, D. B., & Gureckis, T. M. (2014). Is it better to select or to receive? Learning via active and passive hypothesis testing. *Journal of Experimental Psychology: General*, 143, 94–122. doi:10.1037/a0032108
- Martin, E. (1968). Stimulus meaningfulness and paired-associates transfer: An encoding variability hypothesis. *Psychological Review*, 75, 421–441.
- Mata, R., Pachur, T., von Helversen, B., Hertwig, R., Rieskamp, J., & Schooler, L. (2012). Ecological rationality: A framework for understanding and aiding the aging decision maker. *Frontiers in Neuroscience*, 6, 1–19. doi:10.3389/fnins.2012.00019

- Mata, R., von Helversen, B., Karlsson, L., & Cüpper, L. (2012). Adult age differences in categorization and multiple-cue judgment. *Developmental Psychology*, 48, 1188–1201. doi:10.1037/a0026084
- McDaniel, M. A., Cahill, M. J., Robbins, M., & Wiener, C. (2014). Individual differences in learning and transfer: Stable tendencies for learning exemplars versus abstracting rules. *Journal of Experimental Psychology: General*, 143, 668–693. doi:10.1037/a0032963
- McDaniel, M. A., Fadler, C. L., & Pashler, H. (2013). Effects of spaced versus massed training in function learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 1417–1432. doi:10.1037/a0032184
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, 85, 207–238. doi:10.1037/0033-295X.85.3.207
- Meier, B., & Perrig, W. J. (2000). Low reliability of perceptual priming: Consequences for the interpretation of functional dissociations between explicit and implicit memory. *Quarterly Journal of Experimental Psychology: A Human Experimental Psychology*, 53, 211–233. doi:10.1080/713755878
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49–100. doi:10.1006/cogp.1999.0734
- Mogle, J. A., Lovett, B. J., Stawski, R. S., & Sliwinski, M. J. (2008). What’s so special about working memory? An examination of the relationships among working memory, secondary memory, and fluid intelligence. *Psychological Science*, 19, 1071–1077. doi:10.1111/j.1467-9280.2008.02202.x

- Musen, G., & Treisman, A. (1990). Implicit and explicit memory for visual patterns. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 16, 127–137. doi:10.1037/0278-7393.16.1.127
- Muthén, L. K., & Muthén, B. O. (2010). Mplus (Version 6) [Software]. Los Angeles, CA: Muthén & Muthén.
- Newell, B. R., Dunn, J. C., & Kalish, M. L. (2011). Systems of category learning: Fact or fantasy? In B. H. Ross (Ed.), *The psychology of learning & motivation* (Vol. 54, pp. 1–82). San Diego, CA: Academic Press.
- Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 700–708. doi:10.1037/0278-7393.14.4.700
- Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, 101, 53–79. doi:10.1037/0033-295X.101.1.53
- Nosofsky, R. M., & Zaki, S. R. (1998). Dissociations between categorization and recognition in amnesic and normal individuals: An exemplar-based interpretation. *Psychological Science*, 9, 247–255. doi:10.1111/1467-9280.00051
- Oberauer, K. (2009). Design for a working memory. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 51, pp. 45–100). Burlington, MA: Academic Press.
- Olsson, A.-C., Enkvist, T., & Juslin, P. (2006). Go with the flow: How to master a nonlinear multiple-cue judgment task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 1371–1384. doi:10.1037/0278-7393.32.6.1371

- Pachur, T., & Olsson, H. (2012). Type of learning task impacts performance and strategy selection in decision making. *Cognitive Psychology*, 65, 207–240. doi:10.1016/j.cogpsych.2012.03.003
- Palmeri, T. J., & Flanery, M. A. (1999). Learning about categories in the absence of training: Profound amnesia and the relationship between perceptual categorization and recognition memory. *Psychological Science*, 10, 526–530. doi:10.1111/1467-9280.00200
- Palmeri, T. J., & Nosofsky, R. M. (1995). Recognition memory for exceptions to the rule. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 548–568. doi:10.1037/0278-7393.21.3.548
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York, NY: Cambridge University Press.
- Perruchet, P., & Beaveux, P. (1989). Correlational analyses of explicit and implicit memory performance. *Memory & Cognition*, 17, 77–86. doi:10.3758/BF03199559
- Platzer, C., & Bröder, A. (2013). When the rule is ruled out: Exemplars and rules in decisions from memory. *Journal of Behavioral Decision Making*, 26, 429–441. doi:10.1002/bdm.1776
- Raykov, T. (2004). Point and interval estimation of reliability for multiple-component measuring instruments via linear constraint covariance structure modeling. *Structural Equation Modeling*, 11, 342–356.
- Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment*, 28, 164–171. doi:10.1027/1015-5759/a000123

- Rieskamp, J. (2006). Perspectives of probabilistic inferences: Reinforcement learning and an adaptive network compared. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 1355–1370. doi:10.1037/0278-7393.32.6.1355
- Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, 135, 207–236. doi:10.1037/0096-3445.135.2.207
- Roediger, H. L. (1990). Implicit memory: Retention without remembering. *American Psychologist*, 45, 1043–1056. doi:10.1037/0003-066X.45.9.1043
- Rolison, J. J., Evans, J. S. B. T., Walsh, C. R., & Dennis, I. (2011). The role of working memory capacity in multiple-cue probability learning. *Quarterly Journal of Experimental Psychology*, 64, 1494–1514. doi:10.1080/17470218.2011.559586
- Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General*, 126, 211–227. doi:10.1037/0096-3445.126.3.211
- Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object set: The role of surface detail in basic-level object recognition. *Perception*, 33, 217–236.
- Rouder, J. N., & Ratcliff, R. (2004). Comparing categorization models. *Journal of Experimental Psychology: General*, 133, 63–82. doi:10.1037/0096-3445.133.1.63
- Sakamoto, Y., & Love, B. C. (2004). Schematic influences on category learning and recognition memory. *Journal of Experimental Psychology: General*, 133, 534–553. doi:10.1037/0096-3445.133.4.534

- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66, 507–514.
- Schubert, E., & Wolfe, J. (2006). Does timbral brightness scale with frequency and spectral centroid? *Acta Acustica United with Acustica*, 92, 820–825.
- Schunn, C. D., Lovett, M. C., & Reder, L. M. (2001). Awareness and working memory in strategy adaptivity. *Memory & Cognition*, 29, 254–266.  
doi:10.3758/BF03194919
- Shiffrin, R., & Atkinson, R. (1969). Storage and retrieval processes in long-term memory. *Psychological Review*, 76, 179–193.
- Smith, E. E. (2008). The case for implicit category learning. *Cognitive, Affective, & Behavioral Neuroscience*, 8, 3–16. doi:10.3758/CABN.8.1.3
- Smith, E. E., & Grossman, M. (2008). Multiple systems of category learning. *Neuroscience and Biobehavioral Reviews*, 32, 249–264.  
doi:10.1016/j.neubiorev.2007.07.009
- Smith, E. E., Patalano, A. L., & Jonides, J. (1998). Alternative strategies of categorization. *Cognition*, 65, 167–196. doi:10.1016/S0010-0277(97)00043-7
- Speekenbrink, M., & Shanks, D. R. (2010). Learning in a changing environment. *Journal of Experimental Psychology: General*, 139, 266–298.  
doi:10.1037/a0018620
- Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 13515–13522.
- Tomarken, A. J., & Waller, N. G. (2005). Structural equation modeling: Strengths, limitations, and misconceptions. *Annual Review of Clinical Psychology*, 1, 31–65. doi:10.1146/annurev.clinpsy.1.102803.144239

- Tomlinson, T., Marewski, J. N., & Dougherty, M. (2011). Four challenges for cognitive research on the recognition heuristic and a call for a research strategy shift. *Judgment and Decision Making*, 6, 89–99.
- Unsworth, N. (2010). On the division of working memory and long-term memory and their relation to intelligence: A latent variable approach. *Acta Psychologica*, 134, 16–28. doi:10.1016/j.actpsy.2009.11.010
- Unsworth, N., Brewer, G. A., & Spillers, G. J. (2013). Working memory capacity and retrieval from long-term memory: The role of controlled search. *Memory & Cognition*, 41, 242–254. doi:10.3758/s13421-012-0261-x
- Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, 114, 104–132. doi:10.1037/0033-295X.114.1.104
- Unsworth, N., McMillan, B. D., Brewer, G. A., & Spillers, G. J. (2012). Everyday attention failures: An individual differences investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 1765–1772. doi:10.1037/a0028075
- Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory*, 17, 635–654. doi:10.1080/09658210902998047
- von Helversen, B., Karlsson, L., Mata, R., & Wilke, A. (2013). Why does cue polarity information provide benefits in inference problems? The role of strategy selection and knowledge of cue importance. *Acta Psychologica*, 144, 73–82. doi:10.1016/j.actpsy.2013.05.007

von Helversen, B., Mata, R., & Olsson, H. (2010). Do children profit from looking beyond looks? From similarity-based to cue abstraction processes in multiple-cue judgment. *Developmental Psychology*, 46, 220–229.

doi:10.1037/a0016690

von Helversen, B., & Rieskamp, J. (2008). The mapping model: A cognitive theory of quantitative estimation. *Journal of Experimental Psychology: General*, 137, 73–96. doi:10.1037/0096-3445.137.1.73

von Helversen, B., & Rieskamp, J. (2009). Models of quantitative estimations: Rule-based and exemplar-based processes compared. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 867–889.

doi:10.1037/a0015501

Weaver, E. A., & Stewart, T. R. (2012). Dimensions of judgment: Factor analysis of individual differences. *Journal of Behavioral Decision Making*, 25, 402–413.

doi:10.1002/bdm.748

Weber, E. U., Goldstein, W. M., & Barlas, S. (1995). And let us not forget memory: The role of memory processes and techniques in the study of judgment and choice. *The Psychology of Learning and Motivation*, 32, 33–81.

doi:10.1016/S0079-7421(08)60307-2

Wigton, R. S. (1996). Social judgement theory and medical judgement. *Thinking and Reasoning*, 2, 175–190. doi:10.1080/135467896394492

Willingham, D. (1998). A neuropsychological theory of motor skill learning.

*Psychological Review*, 105, 558–584. doi:10.1037/0033-295X.105.3.558

Zaki, S. R., & Kleinschmidt, D. F. (2014). Procedural memory effects in categorization: Evidence for multiple systems or task complexity? *Memory & Cognition*, 42, 508–524. doi:10.3758/s13421-013-0375-9



Zeithamova, D., & Maddox, W. T. (2006). Dual-task interference in perceptual category learning. *Memory & Cognition*, 34, 387–398.

doi:10.3758/BF03193416

Zeithamova, D., & Maddox, W. T. (2007). The role of visuospatial and verbal working memory in perceptual category learning. *Memory & Cognition*, 35, 1380–1398. doi:10.3758/BF03193609

Table 1

*Training and Validation Items Used in the Multiplicative and the Linear Task. The Judgment Criterion Was Derived from Equation 1 (Linear) and Equation 2 (Multiplicative)*

Cue values				Judgment criterion		Set
Cue 1	Cue 2	Cue 3	Cue 4	Multiplicative	Linear	
2	1	0	3	2	14	Training
1	4	1	4	5	22	Training
0	3	1	2	2	13	Training
0	2	3	0	1	12	Training
5	5	4	0	29	43	Training
0	4	5	4	12	26	Training
2	4	3	0	9	26	Training
1	4	3	5	13	27	Training
1	0	2	4	1	12	Training
1	0	0	2	1	6	Training
5	3	3	5	21	40	Training
1	1	5	5	7	22	Training
1	2	0	5	2	15	Training
5	5	0	1	4	36	Training
0	4	3	1	4	19	Training
4	2	1	3	6	27	Training
0	5	2	3	6	22	Training
5	5	2	4	22	43	Training
5	1	3	4	9	33	Training

Cue values				Judgment criterion		Set
Cue 1	Cue 2	Cue 3	Cue 4	Multiplicative	Linear	
4	0	2	4	3	24	Training
1	4	1	5	6	23	Training
3	0	5	5	3	27	Training
0	2	5	0	2	16	Training
1	5	2	4	10	27	Training
3	4	5	5	30	39	Training
3	5	1	4	10	33	Validation
3	4	4	3	21	35	Validation
5	0	3	4	4	30	Validation
3	4	2	5	14	33	Validation
5	0	5	5	4	35	Validation
3	2	0	2	2	20	Validation
2	3	4	0	9	25	Validation
4	5	4	5	36	44	Validation
5	0	5	3	4	33	Validation
4	3	0	1	3	26	Validation
2	1	2	0	3	15	Validation
2	5	2	3	12	30	Validation
4	0	0	2	2	18	Validation
4	1	1	1	4	22	Validation
3	3	3	5	15	32	Validation

Table 2

*Descriptive Statistics for the Memory and Judgment Tasks*

Task	<i>M</i>	<i>SD</i>	Skewness	Kurtosis
Operation span	58.4	11.7	-1.3	2.2
Reading span	57.6	11.8	-1.2	2.1
Symmetry span	29.9	7.4	-0.7	0.1
Recognition (% recalled)	86.5	8.8	-0.7	0.2
Cued recall (% recalled)	41.4	19.6	0.3	-0.2
Free recall (% recalled)	44.6	16.5	0.3	-0.1
Speeded presentation (ms)	55	1023	0.2	5.3
Degraded presentation (ms)	-1293	3471	0.4	3.5
Identification in noise (ms)	-371	788	-0.9	3.1
Linear judgment				
Last training block	6.1	2.4	1.4	4.1
Test (mean)	5.4	1.9	0.5	0.6
Strategy weight ( $n = 262$ )	.22	.28	1.32	0.81
Multiplicative judgment				
Last training block	5.3	1.8	0.6	0.4
Test (mean)	5.1	1.9	0.9	0.5
Strategy weight ( $n = 275$ )	.60	.38	-0.52	-1.26

Table 3

*Model Fits (and Standard Deviations) in the Linear and the Multiplicative Task*

Model fit	Model		
	Baseline	Linear	Exemplar
Linear task			
Training RMSD	9.0 (1.3)	4.5 (1.4)	5.3 (1.6)
Test RMSD	7.3 (1.6)	3.8 (1.4)	5.2 (1.5)
Classification ( <i>n</i> )	17	220	42
Multiplicative task			
Training RMSD	7.3 (1.2)	4.7 (0.9)	4.7 (1.1)
Test RMSD	6.9 (1.9)	4.6 (1.3)	4.2 (1.1)
Classification ( <i>n</i> )	4	99	176

*Note.* RMSD: root mean square deviation